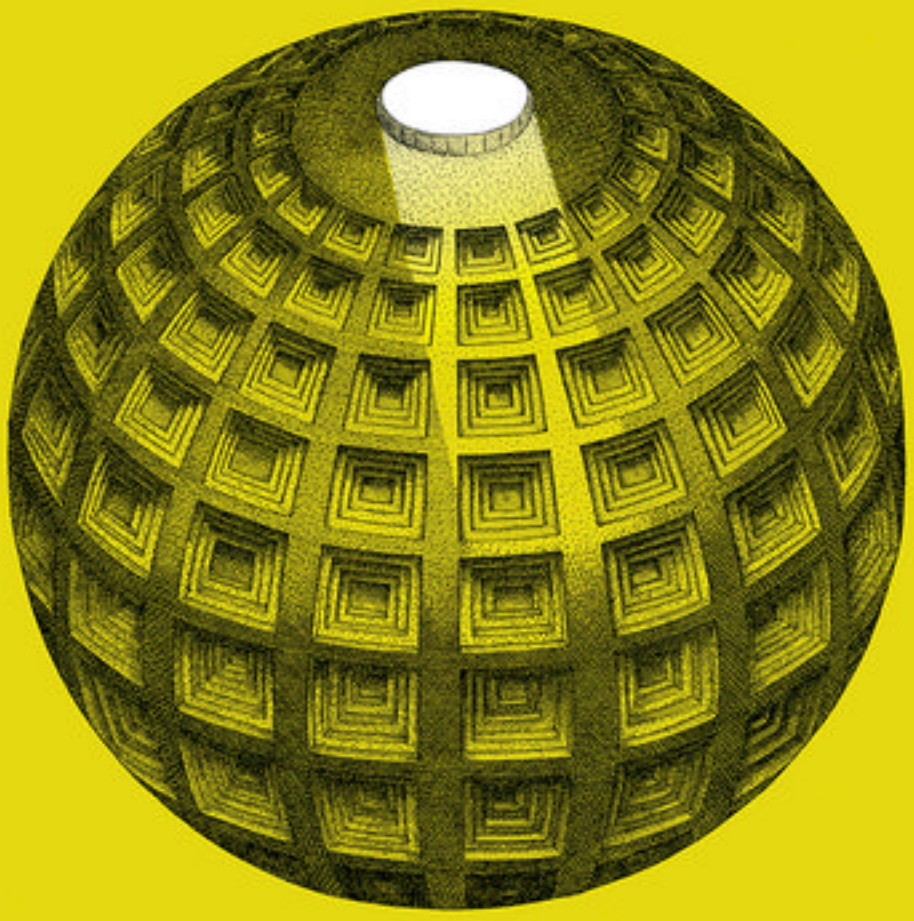
THE OXFORD COMPENDIUM OF VISUAL ILLUSIONS



EDITED BY ARTHUR G. SHAPIRO
AND DEJAN TODOROVIĆ

The Oxford Compendium of Visual Illusions

EDITED BY

Arthur Shapiro
and
Dejan Todorović

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide. Oxford is a registered trade mark of Oxford University Press in the UK and certain other countries.

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America.

© Arthur Shapiro and Dejan Todorović 2017

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, without the prior permission in writing of Oxford University Press, or as expressly permitted by law, by license, or under terms agreed with the appropriate reproduction rights organization. Inquiries concerning reproduction outside the scope of the above should be sent to the Rights Department, Oxford University Press, at the address above.

You must not circulate this work in any other form
and you must impose this same condition on any acquirer.

Library of Congress Cataloging-in-Publication Data
Names: Shapiro, Arthur (Arthur Gilman), editor. | Todorović, Dejan, editor.
Title: The Oxford compendium of visual illusions / edited by Arthur Shapiro and Dejan Todorović.
Other titles: Visual illusions
Description: Oxford : Oxford University Press, [2017] | Includes bibliographical references and index.
Identifiers: LCCN 2015038872 | ISBN 9780199794607 (hardcover : alk. paper)
Subjects: LCSH: Optical illusions. | Visual perception.
Classification: LCC QP495 .O94 2016 | DDC 612.8/4–dc23 LC record available at http://lccn.loc.gov/2015038872

1 3 5 7 9 8 6 4 2
Printed by Sheridan Books, Inc., United States of America

Contents

Contributors xi

How to Use the Website xvii

Introduction xix
Arthur Shapiro and Dejan Todorović

Part I: Introductory General Chapters

1. Early History of Illusions 3
   Nicholas J. Wade
2. Cross-Cultural Studies of Illusions 38
   J. B. Deregowski
3. Visual Illusion in a Comparative Perspective 54
   Kazuo Fujita, Noriyuki Nakamura, and Sota Watanabe
4. An Analysis of Theoretical Approaches to Geometrical-Optical Illusions 64
   Barbara Gillam
5. Visual Illusions in Action 74
   Nicola Bruno
6. Motion Illusions in Man and Machine 79
   Cornelia Fermüller
7. The Visual World as Illusion: The Ones We Know and the Ones We Don't 90
   Stephen Grossberg
8. Visual Illusions? 119
   Jan Koenderink
9. Why the Concept of "Visual Illusions" Is Misleading 139
   Dale Purves, William T. Wojtach, and R. Beau Lotto
10. Where Have All the Illusions Gone? A Critique of the Concept of Illusion 144
    Brian Rogers

Part II: Geometrical

11. Weighted Positional Averaging in the Illusions of the Müller-Lyer Type 159
    Aleksandr Bulatov
12. The Bar-Cross-Ellipse Illusion 165
    Gideon P. Caplovitz, Alex Boswell, and Kyle Killebrew
13. The Spinning Ellipse Speed Illusion 170
    Gideon Paul Caplovitz, Po-Jang Hsieh, Peter J. Kohler, and Katharine B. Porter
14. The Ames Window Illusion and Its Variations 174
    Marcel de Heer and Thomas V. Papathomas
15. Three-Dimensional Müller-Lyer Illusion: Theoretical and Practical Implications 185
    Patricia R. De Lucia
16. Why Do Hills Look So Steep? 190
    Frank H. Durgin and Zhi Li
17. "Shape from Smear": An Illusion of 3D Shape, Made by Finger-Painting With Noise 198
    Roland W. Fleming and Daniel Holtmann-Rice
18. Geometric-Optical Illusions Under Isoluminance? 206
    Kai Hamburger, Thorsten Hansen, and Karl R. Gegenfurtner

19. The Picture Surface Illusion: 3D Biases 2D 209
Sherief Hammad and John M. Kennedy

20. Cast Shadow Illusions 214
Daniel Kersten and Pascal Mamassian

21. The Leaning Tower Illusion 221
Frederick A. A. Kingdom, Ali Yoonessi, and Elena Gheorghiu

22. The Invisible Saddle, or the Cap-or-Cup Illusion 227
Jan Koenderink, Andrea van Doorn, and Johan Wagemans

23. Symmetry and Uprightness in Visually Perceived Forms 234
Lydia M. Maniatis

24. The Bathtub Illusion 238
Lydia M. Maniatis

25. The Pitchroom Illusion: How High Is Up? 241
Leonard Matin, Ethel Matin, Wenxun Li, Todd E. Hudson, and Adam Shavit

26. Geometric Illusions in the Human Face and Body 252
Kazunori Morikawa

27. Dynamic Illusory Size Contrast: Enhanced Relative Size Effects Due to Stimulus Motion 258
Ryan E. B. Mruczek, Christopher D. Blair, Lars Strother, and Gideon P. Caplovitz

28. Size Contrast and Assimilation in the Delboeuf and Ebbinghaus Illusions 262
Ryan E. B. Mruczek, Christopher D. Blair, Lars Strother, and Gideon P. Caplovitz

29. The Occlusion, Configural Shape, and Shrinkage Illusions 269
Stephen E. Palmer and Karen B. Schloss

30. Reverse-Perspective Art and Objects—Illusions in Depth and Motion 275
Thomas V. Papathomas

31. The *New* Moon Illusion 283
Brian Rogers and Stuart Anstis

32. Geometrical Errors Are the Cost of Maintaining the Luminance Contrast Polarity 290
Sergio Roncato

33. Antigravity Slopes: A New Type of Visual Illusion 295
Kokichi Sugihara

34. The Geometrical-Optical Illusions of J. J. Oppel 298
Dejan Todorović

35. Oppel–Kundt Illusion 303
Jiří Wackermann

36. The Shifted-Chessboard Pattern as Paradigm of the Exegesis of Geometrical-Optical Illusions 308
Gerald Westheimer

Part III: Brightness/Lightness/Color

37. A Layered Experience of Lightness and Color 315
Barton L. Anderson

38. Color and Luminance: Afterimages, Combinations, and Flicker 323
Stuart Anstis

39. The White Effect 327
Barbara Blakeslee and Mark E. McCourt

40. The Dungeon Illusion 332
Paola Bressan and Peter Kramer

41. The Contrast Illusion 337
Charles Chubb, Joshua A. Solomon, and George Sperling

42. Illusory Color Spread from Apparent Motion 344
Carol M. Cicerone and Donald D. Hoffman

43. The Reversed Contrast Necker Cube 349
Alessandra Galmonte and Tiziano Agostini

44. Changing the Chevreul Illusion by a Background Luminance Ramp 356
János Geier and Mariann Hudák

45. The Curved Grid Non-Illusions: Eliminating Hermann's Spots and Lingelbach's Scintillation 362
János Geier and Mariann Hudák

46. The Staircase Gelb Illusion 367
Alan Gilchrist

47. The Breathing Light Illusion: Illusory Size and Brightness Variation Induced by Motion 373
Simone Gori, Enrico Giora, and D. Alan Stubbs

48. Large Shift in Brightness Induced by Motion in Context 379
Sang Wook Hong

49. The Chromatic Mach Card 382
Anya Hurlbert

50. Color Assimilation 388
Frederick A. A. Kingdom

51. When Light Looks Like Paint 393
Frederick A. A. Kingdom

52. The Scintillating Grid 400
Bernd Lingelbach

53. Second-Order Mach Bands, Chevreul, and Craik-O'Brien-Cornsweet Illusions 404
Zhong-Lin Lu and George Sperling

54. Vasarely's Nested Squares and the Alternating Brightness Star Illusion 407
Susana Martinez-Conde and Stephen L. Macknik

55. Grating Induction 415
Mark E. McCourt and Barbara Blakeslee

56. Illusory Variations in Apparent Dot Brightness Induced by Density Modulations 422
Jeffrey B. Mulligan

57. On the Watercolor Illusion 426
Baingio Pinna

58. The Chinese Lantern Illusion 437
Sergio Roncato, Sandro Bettella, and Clara Casco

59. The Wedding Cake Illusion: Interaction of Geometric and Photometric Factors in Induced Contrast and Assimilation 442
Branka Spehar and Colin W. G. Clifford

60. Filling-In Between Contours 448
Rob van Lier

61. The Glare Effect 454
Daniele Zavagno and Olga Daneyko

Part IV: Motion Based

62. Improbable Illusory Contours 463
Barton L. Anderson

63. Low-Level Motion Illusions 467
Stuart Anstis

64. High-Level Organization of Motion: Ambiguous, Primed, Sliding, and Flashed 475
Stuart Anstis

65. Backscroll Illusion 486
Kiyoshi Fujimoto

66. The Rotating Tilted Lines Illusion: Rotation Instead of Expansion—A Peculiar Case of Motion Misperception 489
Simone Gori

67. The Enigmatic Enigma Illusion 496
Kai Hamburger

68. The Fraser-Wilcox Illusion and Its Extension 500
Akiyoshi Kitaoka

69. Induced Motion 512
Jasmin Léveillé and Arash Yazdanbakhsh

70. The Freezing Rotation Illusion 516
Erika N. Lorincz and Max R. Dürsteler

71. Second-Order Reversed Phi 522
Zhong-Lin Lu and George Sperling

72. Attention-Generated Apparent Motion 527
Zhong-Lin Lu and George Sperling

73. Two-Stroke Apparent Motion 531
George Mather

74. On the Pinna Illusion 536
Baingio Pinna

75. Color Wagon-Wheel Illusion 548
Arthur G. Shapiro and William Kistler

76. The Aperture Problem: Illusions Arising During the Integration and Segmentation of Motion Within and Across Apertures 556
Maggie Shiffrar

77. Paths of Apparent Human Motion Follow Motor Constraints 563
Maggie Shiffrar and Christina Joseph

78. The Motion Standstill Illusion 569
George Sperling, Son-Hee Lyu, Chia-Huei Tseng, and Zhong-Lin Lu

79. Objectless Motion: The Pedestalled Motion Paradigm 573
George Sperling and Zhong-Lin Lu

80. Silencing the Awareness of Change 577
Jordan W. Suchow and George A. Alvarez

81. The Kayahara Silhouette Illusion 582
Nikolaus F. Troje

82. The Motion Aftereffect 586
Frans A. J. Verstraten and Peter J. Bex

83. High Phi and Ghost Phi: Extreme Motion Illusions 592
Mark Wexler

84. Stereokinetic Phenomena 597
Mario Zanforlin

85. Motion Illusions in Static Patterns 601
Johannes M. Zanker

Part V: Faces

86. The Venus Effect 609
Marco Bertamini and Richard Latto

87. The Hollow-Mask Illusion and Variations 614
Thomas V. Papathomas

88. The Illusion of Sex 620
Richard Russell

89. The Bogart Effect 624
Sharon Gilad-Gutnick, Rohan Varma, and Pawan Sinha

90. The Presidential Illusion 628
Sharon Gilad-Gutnick and Pawan Sinha

91. About Face: The Margaret Thatcher Illusion 633
Peter Thompson

92. The Mona Lisa Gaze Effect 639
Dejan Todorović

93. The Wobbling Face Illusion 648
Sayako Ueda and Akiyoshi Kitaoka

94. Adaptation Aftereffects in the Perception of Faces 655
Michael A. Webster

Part VI: Grouping and Organization

95. Ambiguous Figures Moving Forward 663
Lori J. Bernstein

96. The Scramble Illusion: Texture Metamers 668
Charles Chubb, Joseph Darcy, Michael S. Landy, John Econopouly, Jong-Ho Nam, Dan Bindman, and George Sperling

97. Amodally Completed Angles 673
Walter Gerbino

98. Subjective Contours 679
Barbara Gillam

99. The Ternus Effect 685
Elisabeth Hein

100. Two Sinusoids: 6 – 1 Perceptions 692
Jan Kremlacek

101. The Illusions of Numerosity 696
Riccardo Luccio

102. The Aperture Capture Illusion 700
Evan M. Palmer and Philip J. Kellman

Part VII: Attention

103. Motion-Induced Blindness 709
Yoram Bonneh

104. Inattentional Blindness and the *Illusion* of Attention 715
Daniel J. Simons

Part VIII: Binocular Vision/Stereopsis

105. Binocular Rivalry: The Illusion of Disappearance 721
Randolph Blake

106. Stereo Rotation Standstill and Related Illusions 726
Max R. Dürsteler and Erika N. Lorincz

107. The Graph-Paper Effect: A Moving, Illusory, Stereoscopic Texture 733
Mark Georgeson

Part IX: Adaptation

108. Adaptation to Brightness Change, Contours, Jogging, and Apparent Motion 745
Stuart Anstis

109. The Color Dove Illusion: Chromatic Filling-In Effect Following a Spatial-Temporal Edge 752
Yuval Barkan and Hedva Spitzer

110. Blur Adaptation and Induction 756
Michael A. Webster

Part X: Conflicting Scale or Other Information

111. Hybrid Image Illusion 763
Aude Oliva and Philippe G. Schyns

112. Contrast Asynchronies 767
Arthur G. Shapiro

113. Hidden Images 774
Nicholas J. Wade

Part XI: Multisensory

114. The Three-Dimensional Necker Cube 783
Nicola Bruno

115. The McGurk Effect and the Primacy of Multisensory Perception 787
James W. Dias, Theresa C. Cook, and Lawrence D. Rosenblum

Index 793

Contributors

Tiziano Agostini
Department of Life Sciences
University of Trieste
Trieste, Italy

George A. Alvarez
Department of Psychology
Harvard University
Cambridge, MA, USA

Barton L. Anderson
University of Sydney
School of Psychology
Sydney, NSW, Australia

Stuart Anstis
Department of Psychology
University of California, San Diego
La Jolla, CA, USA

Yuval Barkan
Department of Biomedical Engineering
Tel Aviv University
Tel Aviv, Israel

R. Beau Lotto
Department of Computer Science
Goldsmiths, University of London
London, UK

Lori J. Bernstein
Department of Supportive Care
Princess Margaret Cancer Centre;
Department of Psychiatry
University of Toronto
Toronto, ON, Canada

Marco Bertamini
Department of Psychological Sciences
University of Liverpool
Liverpool, UK

Sandro Bettella
Dipartimento di Psicologia Generale
Università di Padova
Padova, Italy

Peter J. Bex
Department of Psychology
Northeastern University
Boston, MA, USA

Dan Bindman
Aleks Corporation
Irvine, CA, USA

Christopher D. Blair
Department of Psychology
University of Nevada, Reno
Reno, NV, USA

Randolph Blake
Department of Psychology
Vanderbilt University
Nashville, TN, USA;
Department of Brain and Cognitive Science
Seoul National University
Gwanak, Seoul, Korea

Barbara Blakeslee
Center for Visual and Cognitive Neuroscience
Department of Psychology
North Dakota State University
Fargo, ND, USA

Yoram Bonneh
Department of Optometry and Vision Science
Bar-Ilan University
Ramat-Gan, Israel

Alex Boswell
Department of Psychology
University of Nevada, Reno
Reno, NV, USA

Paola Bressan
Dipartimento di Psicologia Generale
Università di Padova
Padova, Italy

Nicola Bruno
Università di Parma
Dipartimento di Neuroscienze
Parma, Italy

Aleksandr Bulatov
Institute of Biological Systems and Genetics Research
Lithuanian University of Health Sciences
Kaunas, Lithuania

Gideon P. Caplovitz
Department of Psychology
University of Nevada, Reno
Reno, NV, USA

Clara Casco
Dipartimento di Psicologia Generale
Università di Padova
Padova, Italy

Charles Chubb
Department of Cognitive Sciences
University of California
Irvine, CA, USA

Carol M. Cicerone
Department of Cognitive Sciences
University of California, Irvine
Irvine, CA, USA

Colin W. G. Clifford
School of Psychology
University of New South Wales
Sydney, NSW, Australia

Theresa C. Cook
Department of Psychology
University of California, Riverside
Riverside, CA, USA

Olga Daneyko
Psychology Department
University of Milano-Bicocca
Milano, Italy

Joseph Darcy
Oracle Corporation
San Francisco, CA, USA

Patricia R. DeLucia
Department of Psychological Sciences
Texas Tech University
Lubbock, TX, USA

J. B. Deręgowski
Department of Psychology
University of Aberdeen
Aberdeen, Scotland, UK

James W. Dias
Department of Psychology
University of California, Riverside
Riverside, CA, USA

Frank H. Durgin
Department of Psychology
Swarthmore College
Swarthmore, PA, USA

Max R. Dürsteler
Department of Neurology
University Hospital Zurich
Zürich, Switzerland

John Economopouly
Logic Design
Nicosia, Cyprus

Cornelia Fermüller
Computer Vision Laboratory
Institute for Advanced Computer Studies
University of Maryland
College Park, MD, USA

Roland W. Fleming
Department of Psychology
Justus Liebig University
Giessen, Germany

Kiyoshi Fujimoto
Subaru Europe
Zaventem, Belgium

Kazuo Fujita
Graduate School of Letters
Kyoto University
Kyoto, Japan

Alessandra Galmonte
Department of Neurosciences, Biomedicine and Movement Sciences
University of Verona
Verona, Italy

Karl R. Gegenfurtner
Department of Psychology
Justus Liebig University
Giessen, Germany

János Geier
Stereo Vision Ltd.
Budapest, Hungary

Mark Georgeson
School of Life & Health Sciences
Aston University
Birmingham, UK

Walter Gerbino
Department of Life Sciences
University of Trieste
Trieste, Italy

Elena Gheorghiu
Department of Psychology
University of Stirling
Stirling, Scotland, UK

Sharon Gilad-Gutnick
Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology
Cambridge, MA, USA

Alan Gilchrist
Department of Psychology
Rutgers University
Newark, NJ, USA

Barbara Gillam
School of Psychology
University of New South Wales
Sydney, NSW, Australia

Enrico Giora
Faculty of Psychology
University "Vita-Salute" San Raffaele
Milan, Italy

Simone Gori
Department of Human and Social Science
University of Bergamo
Bergamo, Italy;
IRCCS E. Medea
Bosisio Parini, Lecco, Italy

Stephen Grossberg
Center for Adaptive Systems, Graduate Program in Cognitive and Neural Systems
Departments of Mathematics, Psychological & Brain Sciences, and Biomedical Engineering
Boston University
Boston, MA, USA

Kai Hamburger
Department of Psychology
Justus Liebig University
Giessen, Germany

Sherief Hammad
Department of Psychology
University of Toronto at Scarborough
Toronto, ON, Canada

Thorsten Hansen
Department of Psychology
Justus Liebig University
Giessen, Germany

Marcel de Heer
Nieuwezijds Voorburgwal 173
Amsterdam, The Netherlands

Elisabeth Hein
Department of Psychology
University of Tübingen
Tübingen, Germany

Donald D. Hoffman
Department of Cognitive Sciences
University of California, Irvine
Irvine, CA, USA

Daniel Holtmann-Rice
Department of Computer Science
Yale University
New Haven, CT, USA

Sang Wook Hong
Department of Psychology
Florida Atlantic University
Boca Raton, FL, USA

Po-Jang Hsieh
Neuroscience & Behavioral Disorders Program
Graduate Medical School
Duke-National University of Singapore
Singapore

Mariann Hudák
Department of General Psychology
Pázmány Péter Catholic University
Budapest, Hungary

Todd E. Hudson
Department of Psychology
New York University
New York, NY, USA

Anya Hurlbert
Institute of Neuroscience
Newcastle University
Newcastle upon Tyne, UK

Christina Joseph
Department of Psychology
Rutgers University
Newark, NJ, USA

Philip J. Kellman
Department of Psychology
University of California, Los Angeles
Los Angeles, CA, USA

John M. Kennedy
Department of Psychology
University of Toronto at Scarborough
Toronto, ON, Canada

Daniel Kersten
Department of Psychology
University of Minnesota
Minneapolis, MN, USA;
Department of Brain and Cognitive Engineering
Korea University
Seoul, South Korea

Kyle Killebrew
Department of Psychology
University of Nevada, Reno
Reno, NV, USA

Frederick A. A. Kingdom
McGill Vision Research
Department of Ophthalmology
Montréal, QC, Canada

William Kistler
Department of Psychology
American University
Washington, DC, USA

Akiyoshi Kitaoka
Department of Psychology
Ritsumeikan University
Kyoto, Japan

Jan Koenderink
Laboratory of Experimental Psychology
University of Leuven
Leuven, Belgium;
Psychological Laboratory
Utrecht University
Utrecht, The Netherlands

Peter J. Kohler
Psychological and Brain Sciences
Dartmouth College
Dartmouth, NH, USA

Peter Kramer
Dipartimento di Psicologia Generale
University of Padua
Padova, Italy

Jan Kremláček
Faculty of Medicine in Hradec Kralove
Charles University
Prague, Czech Republic

Michael S. Landy
Department of Psychology and Center for Neural Science
New York University
New York, NY, USA

Richard Latto
Department of Psychology
University of Liverpool
Liverpool, UK

Jasmin Léveillé
Scientific Systems Company Inc.
Woburn, MA, USA

Wenxun Li
Clarence H. Graham Memorial Laboratory of Visual Science
Department of Psychology
Columbia University
New York, NY, USA

Zhi Li
Department of Psychology and Behavioral Sciences
Zhejiang University
Hangzhou, China

Bernd Lingelbach
The Barn—Optical Phenomena
Abtsgmuend, Germany

Erika Nora Lorincz
Geneva University Hospitals Psychiatry Unit of Mental Development
Department of Mental Health and Psychiatry
Geneva, Switzerland

Zhong-Lin Lu
Laboratory of Brain Processes (Lobes)
Center for Cognitive and Brain Sciences
Center for Cognitive and Behavioral Brain Imaging
Department of Psychology
The Ohio State University
Columbus, OH, USA

Riccardo Luccio
Department of Life Sciences
University of Trieste
Trieste, Italy

Son-Hee Lyu
Department of Cognitive Sciences
University of California
Irvine, CA, USA

Stephen L. Macknik
State University of New York (SUNY) Downstate Medical Center
Brooklyn, NY, USA

Pascal Mamassian
Laboratoire des Systèmes Perceptifs
Ecole Normale Supérieure
Paris, France

Lydia M. Maniatis
Profitou Ilia 4
Athens, Greece

Susana Martinez-Conde
State University of New York (SUNY) Downstate Medical Center
Brooklyn, NY, USA

George Mather
School of Psychology
University of Lincoln
Lincoln, UK

Ethel Matin
Department of Psychology
Long Island University
Brookville, NY, USA

Leonard Matin
Clarence H. Graham Memorial Laboratory of Visual Science
Department of Psychology
Columbia University
New York, NY, USA

Mark E. McCourt
Center for Visual and Cognitive Neuroscience
Department of Psychology
North Dakota State University
Fargo, ND, USA

Kazunori Morikawa
School of Human Sciences
Osaka University
Osaka, Japan

Ryan E. B. Mruczek
Department of Psychology
Worcester State University
Worcester, MA, USA

Jeffrey B. Mulligan
Human Systems Integration Division
NASA Ames Research Center
Mountain View, CA, USA

Noriyuki Nakamura
Faculty of Human Sciences
Toyo Gakuen University
Tokyo, Japan

Jong-Ho Nam
Department of Psychology
The Catholic University of Korea
Puchon City, Korea

Aude Oliva
Computer Science and Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge, MA, USA

Evan M. Palmer
Department of Psychology
Wichita State University
Wichita, KS, USA

Stephen E. Palmer
Department of Psychology
University of California, Berkeley
Berkeley, CA, USA

Thomas V. Papathomas
Department of Biomedical Engineering and Laboratory of Vision Research
Center for Cognitive Science
Rutgers University
New Brunswick, NJ, USA

Baingio Pinna
Department of Humanity and Social Sciences
University of Sassari
Sassari, Italy

Katharine B. Porter
Department of Psychology
Harvard University
Cambridge, MA, USA

Dale Purves
Department of Neurobiology
Duke University
Durham, NC, USA;
Neuroscience and Behavioral Disorders Program
Duke-National University of Singapore
Singapore

Brian Rogers
Department of Experimental Psychology
University of Oxford
Oxford, UK;
Department of General Psychology
St. Petersburg State University
St. Petersburg, Russia

Sergio Roncato
Dipartimento di Psicologia Generale
Università di Padova
Padova, Italy

Lawrence D. Rosenblum
Department of Psychology
University of California, Riverside
Riverside, CA, USA

Richard Russell
Psychology Department
Gettysburg College
Gettysburg, PA, USA

Karen B. Schloss
Department of Psychology
University of Wisconsin
Madison, WI, USA

Philippe G. Schyns
Institute of Neuroscience and Psychology
University of Glasgow
Glasgow, Scotland, UK

Arthur Shapiro
Departments of Psychology and Computer Science
American University
Washington, DC, USA

Adam Shavit
Psychology Department
Hunter College of the City University of New York
New York, NY, USA

Maggie Shiffrar
Department of Psychology
California State University
Northridge, CA, USA

Daniel J. Simons
Department of Psychology
University of Illinois
Champaign, IL, USA

Pawan Sinha
Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology
Cambridge, MA, USA

Joshua A. Solomon
Division of Optometry and Visual Science
City University
London, UK

Branka Spehar
School of Psychology
University of New South Wales
Sydney, NSW, Australia

George Sperling
Department of Cognitive Sciences
Department of Neurobiology and Behavior
University of California, Irvine
Irvine, CA, USA

Hedva Spitzer
Department of Biomedical Engineering
Tel Aviv University
Tel Aviv, Israel

Lars Strother
Department of Psychology
University of Nevada, Reno
Reno, NV, USA

D. Alan Stubbs
Department of Psychology
University of Maine
Orono, ME, USA

Jordan W. Suchow
Department of Psychology
Harvard University
Cambridge, MA, USA

Kokichi Sugihara
Meiji Institute for Advanced Study of Mathematical Sciences
Meiji University
Tokyo, Japan

Peter Thompson
Department of Psychology
University of York
York, UK

Dejan Todorović
Laboratory of Experimental Psychology
Department of Psychology
University of Belgrade
Belgrade, Serbia

Nikolaus F. Troje
Department of Psychology
Queen's University
Kingston, ON, Canada

Chia-Huei Tseng
Research Institute of Electrical Communication
Tohoku University
Sendai, Japan

Sayako Ueda
RIKEN BSI-TOYOTA Collaboration Center
RIKEN Brain Science Institute
Saitama, Japan

Andrea van Doorn
Laboratory of Experimental Psychology
University of Leuven
Leuven, Belgium;
Psychological Laboratory
Utrecht University
Utrecht, The Netherlands

Rob van Lier
Donders Institute for Brain, Cognition and Behaviour
Centre for Cognition
Radboud University
Nijmegen, The Netherlands

Rohan Varma
Leland High School
San Jose, CA, USA

Frans A. J. Verstraten
School of Psychology
University of Sydney
Sydney, NSW, Australia

Jiří Wackermann
Neuroscience Consulting & Solutions
Gutach i.Br., Germany

Nicholas J. Wade
Psychology
University of Dundee
Dundee, Scotland, UK

Johan Wagemans
Laboratory of Experimental Psychology
University of Leuven
Leuven, Belgium

Sota Watanabe
Course for School Teachers
Osaka Kyoiku University
Osaka, Japan

Michael A. Webster
Department of Psychology
University of Nevada, Reno
Reno, NV, USA

Gerald Westheimer
Division of Neurobiology
University of California, Berkeley
Berkeley, CA, USA

Mark Wexler
Laboratoire Psychologie de la Perception
Centre National de la Recherche Scientifique
Université Paris Descartes
Paris, France

William T. Wojtach
Center for Cognitive Neuroscience
Duke University
Durham, NC, USA

Arash Yazdanbakhsh
Vision Laboratory
Department of Psychological & Brain Sciences
Center for Systems Neuroscience (CSN)
Graduate Program for Neuroscience
Boston University
Boston, MA, USA

Ali Yoonessi
School of Advanced Medical Technologies
Tehran University of Medical Sciences
Tehran, Iran

Mario Zanforlin
Department of Psychology
Università di Padova
Padua, Italy

Johannes M. Zanker
Department of Psychology
Royal Holloway University of London
London, UK

Daniele Zavagno
Psychology Department
University of Milano-Bicocca
Milano, Italy

Introduction

Arthur Shapiro and Dejan Todorović

> Strictly speaking, the concept of illusion has no place in psychology because no experience actually copies "reality."
>
> —Edwin G. Boring,
> *Sensation and Perception in the History of Psychology*, 1942

This quotation draws attention to one of the major paradoxes of human experience: when we open our eyes and see a world full of objects, everything about the world appears "real," and we act as if what we see before us is, in fact, reality. However, a moment's contemplation tells us that our perception is—at best—a limited representation of the "real" world since we are completely oblivious to much that surrounds us. We cannot perceive what we cannot sense. Our sensory palette will not allow us to perceive radio waves, x-rays, ultrasonic pressure waves like bat sonar, odorless and colorless chemicals, and so on. Some aspects of the world change too fast for us to perceive: for example, the flickering of a fluorescent light and the motion of a bullet. Other aspects change too slowly: the motion of the hour hand on the clock and the stars in the sky. The microscopic world (bacteria, viruses, cells, DNA, etc.) is too small for us to see directly and yet has proved to be at least as important for human history as those aspects of the world that our visual system evolved to experience.

If, as stated in the epigraph, "no experience actually copies 'reality,'" what, then, are we experiencing? Apparently, the brain constructs a perceptual theater from the very narrow slice of the real world to which our senses are capable of responding. The perceptual theater is wonderfully functional and all-encompassing; we experience a seamless three-dimensional integration of visual, auditory, and tactile information. The objects in the world appear to have qualities such as color, weight, and texture, and when we move within this world, our actions and the actions of objects seem to follow predictable rules. Our perceptions can warn us of danger, teach us about the location of food, give us pleasure, and allow us to interact with others who seem to share similar conceptions of the world. Indeed, our perceptual theater is so seamless and so lawful that it is hard to believe that our perception is only a representation of reality and not reality itself.

However, every so often, there is a glitch in the machinery that runs our perceptual theater—something unusual happens that tells us that what we perceive does not match our previous conceptions of reality—or there is an odd conflict that makes us question the information presented by our senses.

For example, after noting the mid-gray rectangle in Figure 1a, consider Figure 1b, in which the same rectangle is now divided in half by a column with white triangles on the left and black triangles on the right. The column makes the rectangle appear inhomogeneous: darker to the left of the column and brighter to the right. The effect can be enhanced by adding more columns. In Figure 1c, which contains two columns of black and white triangles (mirror images of each other), the area between the columns in the middle portion of the rectangle now looks lighter than the two flanking sides. In Figure 1d, with five columns of triangles dividing the rectangle into six portions, the dominant perception is now of adjoining light and dark (upright and inverted) fir trees (see Flynn & Shapiro, 2014; Todorović, 2014). The perceived change in appearance is even more evident in a dynamic version of the display in which one of the columns moves back and forth (as can be seen in the movie on the website). The images presented in Figure 1b–d and the interactive movie therefore create a conflict between our perception and a naive intuition that we perceive an exact representation of the world. Such phenomena require us to think about factors that shape our perception and, at the very least, make us realize that what we see is not as obvious as our intuitions might lead us to believe.

Phenomena that generate a conflict between perception and reality (or perception and our conception of reality) are, for better or worse, referred to as "illusions." An excellent history of illusions can be found in chapter 1 by Nicholas Wade. In this introduction, we remind readers that the study of illusions has an extensive history. For instance, the moon illusion (the fact that the moon looks larger at horizon than at zenith) seems to be referred to in 6th-century bce Mesopotamian clay tablets; induced

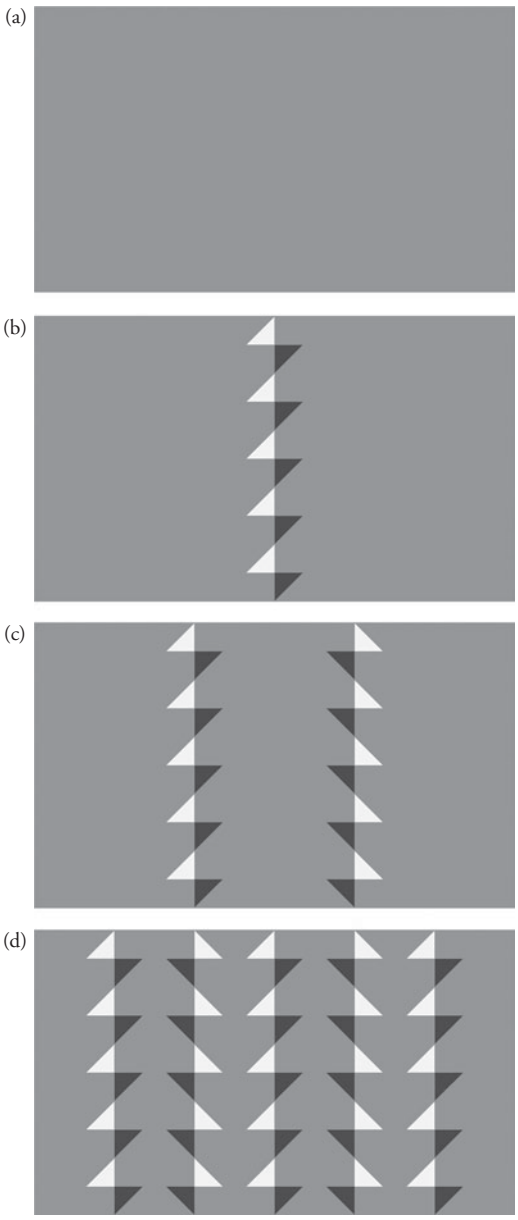


Figure 1. Example of a lightness illusion. (Authors' representation).

motion (the illusory sense that an object or we ourselves are moving when in fact another object nearby moves) was discussed by Euclid (3rd century bce) and Lucretius (1st century bce); color contrast (the fact that our perception of a color depends on other colors adjacent to it, as in Figure 1b–d) was known to Aristotle (4th century bce) and Ibn al-Haytham (11th century ce).

The modern period of the study of illusions arguably started with J. J. Oppel's (1854–1855) paper on what he called geometrical-optical illusions—a term used to this day. Following Oppel's work, about 130 to 140 illusion-related publications appeared in the 19th century (estimate from Zusne, 1968), and research expanded after the publication of the Müller-Lyer illusion in 1889. Ebbinghaus (1913) noted that in the 1890s alone there were "more than 40 independent, occasionally voluminous papers on [geometric illusions]"; among these papers were a number of informative overviews of phenomena, research, and theories, such as Sanford's (1898), Pierce's (1898), and Wundt's (1898).

The growth of the academic study of illusions continued unabated in the 20th century. One estimate places the number of 20th-century publications on illusions to be greater than 1,100 by 1986 (Zusne, 1986); a more recent tally does not seem to be available. Academic book-length treatments of illusions in the past 100 years include Luckiesh (1922), Tolansky (1964), Robinson (1972/1998), Coren and Girgus (1978), Wade (1982, 2006), Fineman (1996), Da Pos and Zambianchi (1996), Ninio (2001), Goto and Tanaka (2005), Gregory (2009), and Vicario (2011). Near the end of the 20th century, illusions became the subject of many popular publications and Internet sites, some created by scientists. Of particular note is Michael Bach's website (http://www.michaelbach.de/ot/), which presents many interactive illusions along with reliable historical references from vision research literature.

PURPOSE OF THE COMPENDIUM

This book is a collection of chapters about visual illusions, written by the researchers who have created or conducted research on the illusions. The desire for such a volume arose because, over the past 30 years, the vision research literature has seen a sharp increase in the number of published visual phenomena; some are new discoveries, and some are rediscoveries, albeit in a different and instructive form. The increase in the number of published illusions is most likely due to the rapid advances in computing and display technology (and also to the increase in the number of vision scientists). The last quarter of the 20th century may be heralded as the onset of the computer era of illusions, which can be differentiated from the paper, ink, and electric light illusions of the 19th century and most of the 20th by their reliance on digital and moving images.

One of the current ways to examine changes in the mentions of illusions over time is with Ngrams, a free service provided by Google that tracks the number of times particular words have appeared in print in Google's database of English-language books. Figure 2 plots the terms "motion

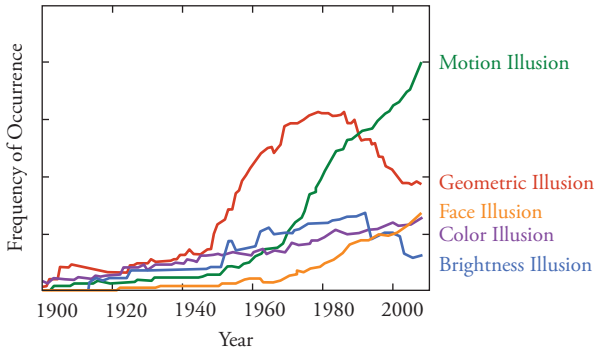


Figure 2. Graph of frequency of several illusion-related concepts over time, based on Google Ngram data. (Authors' graph based on Google Books Ngram viewer http://books.google.com/ngrams.)

illusion," "face illusion," "color illusion," "geometric illusion," and "brightness illusion" from the years 1900 to 2008. Consistent with the idea that technology is transforming the types of phenomena that are being created and studied, the terms for the types of illusions that substantially benefit from advances in computer technology ("motion illusion," "face illusion," and, to a lesser extent, "color illusion") have been increasing in frequency of appearance since the latter half of the 20th century, whereas illusions that could be created with drawings or printed paper ("geometric illusion" and "brightness illusion") peaked in the 1980s and 1990s.

The biggest change, however, seems to be to the speed of dissemination of illusions. Currently, websites, blogs, and other social media are dedicated to the distribution of visual phenomena. While many of these sites link to or discuss the original research, other sites often post, share, or present these illusions for entertainment or for other reasons quite apart from the original purposes of the research. The proliferation accentuates a longstanding issue for artists and authors in all domains: How does the creator of a piece of artwork maintain some semblance of the original intent once the artwork is out of the originator's control? The issue of author intent is perhaps more keen with visual illusions since many phenomena are created by scientists as a result of their research into visual perception and the brain, but illusions are often repackaged to attract attention on the Internet. Such public use of the illusions should be expected (and many times encouraged); the phenomena prompt people to appreciate some of the complexities of the brain's interaction with the external environment. One of the goals of this text is to give those interested in the illusions an authoritative place to which to trace them—to either the originator or at least to a researcher who has worked on the illusion.

DEFINITION AND CLASSIFICATION OF ILLUSIONS

When organizing a collection of chapters such as in this volume, we as editors are faced with two age-old problems: definition and classification.

As stated earlier, illusions are often defined in terms of some discrepancy between reality and perception. One problem with this definition (as expressed by Boring in the epigraph) is that if "no experience actually copies 'reality,'" then everything we experience is an illusion. Even if we do not go that far, the usual definition is certainly too broad, because many cases of perceptual discrepancies from reality are usually not classified as illusions. For example, is it an illusion when we see depth where there is no real depth, such as in realistic paintings, photographs, mirrors, or stereograms? Is all perception of color illusory, given Newton's observation that "the rays, to speak properly, are not coloured"? (Newton, 1730, p. 125) Furthermore, the many cases where we cannot sense physical energy or an event (radio-waves, x-rays, too fast or too slow motions, etc.) are certainly discrepancies between reality and perception.

Such examples illustrate that it is not easy to find a clean definition of the term "illusion" that would not be so general as to render the term meaningless. However, in spite of these semantic problems, vision scientists tend to know an illusion when they see one, even if they may not be able to define the term in a satisfactory manner. Certainly, many classical, paradigmatic examples are generally labeled as illusions, and new examples are steadily generated, as witnessed by the papers in this compendium.

While the definition of "illusion" is slippery, classifying illusions is even more problematic. Vicario (2011) lists 26 historical classification schemes, including those by Wundt, Ebbinghaus, Luckiesh, Robinson, and Gregory. When faced with this problem, we are reminded of J. L. Borges's (1964) classification of animals in the fictitious Chinese *Celestial Emporium of Benevolent Knowledge*: "(a) belonging to the emperor, (b) embalmed, (c) tame, (d) sucking pigs, (e) sirens, (f) fabulous, (g) stray dogs, (h) included in the present classification, (i) frenzied, (j) innumerable, (k) drawn with a very fine camelhair brush, (l) et cetera, (m) having just broken the water pitcher, (n) that from a long way off look like flies." Can we offer in this compendium a categorical scheme that is logically superior to the scheme in Borges's *Emporium*? In a word, unlikely.

There is no generally accepted, useful, comprehensive, and logically satisfactory classification of illusions. However, perhaps the simplest and cleanest is the one by Thiéry (1895), who divided illusions according to "object of study": illusions of extent, color, and motion. We have incorporated this idea into this compendium by including sections dealing with geometry, lightness/brightness/color, and motion. However, in a somewhat Borgesian move of breaking the already established principle of classification, we have also included a number of sections that logically overlap with the topics of the first three sections but still represent coherent domains of study. These are sections on faces, grouping and organization, attention, stereo, adaptation, conflicting scale or other information, and multisensory effects. This classification was driven in part by the range of submissions we received and seems to reflect the current interest of researchers in what can very tentatively be called "the illusion community." Preceding these sections is an introductory section that contains overviews of wider topics relevant to the study of illusions.

THEORETICAL APPROACHES TO ILLUSIONS

Most of the chapters in the compendium present a theoretical explanation of the illusions; however, authors often differ in their theoretical starting points, and even those with similar foundations often present different formulations for their explanations. Readers not acquainted with vision science (and even many who are) may find this situation confusing; we therefore say a few words here about the theoretical challenges one faces when trying to explain and understand visual phenomena.

Explanations of visual phenomena often start from foundational principles based on (a) the physics of light, (b) the physiology of our sensory responses, (c) computational models about how the brain processes sensory information, or (d) psychological or behavioral principles that govern our perceptual interpretation of the world. Each of the levels of analysis adds different insights to most illusions; in general, though, different explanations create

longstanding disagreements between researchers concerning the underlying cause(s) of an illusion. For instance, Isaac Newton's and James Clerk Maxwell's interest in the phenomena of color metamerism led to deep understanding of the trichromacy of color vision; however, even with the great success of Newton's and Maxwell's explanations, vision scientists had fierce arguments concerning the correct level of explanation for color phenomena (for review, see Turner, 1994/2014).

The fractious nature of the disagreement among vision researchers was captured by Ebbinghaus (1913) over a century ago:

> When one shows such an illusory pattern to an intelligent observer, then he very soon finds, as it is occasionally stressed with satisfaction and self-confidence in published reports, an appropriate and immediately obvious explanation. Such a clear issue can only be accounted for in such and such manner, otherwise one would deny one's very capacity of theoretical thinking. The solution is confirmed by some easily found evidence, although something so obvious isn't even in need of any proof.
>
> However, he is astonished to find that another observer not only has not found the correct solution, due to his lower intelligence, but even, when he hears it, maliciously rejects it, and insists on a different solution with the same self-assuredness and the same claim of obviousness. And if he asks a third or fourth person, the same experience is repeated: his own explanation, immediately self-evident to him, is met with opposition and counter-argumentation, and he has to endure other accounts, which appear to him completely misguided. So everyone . . . pursues the matter in accord with his own ideas. (pp. 53–54)

Ebbinghaus says that his story "may be instructive about psychology," and indeed, a century later the anecdote is humorously familiar to researchers who have tried to convince others of the centrality of a particular explanation—and, to be fair, the story resonates with research in many empirical endeavors, not just the study of visual phenomena.

But perhaps the most telling aspect of the anecdote is the conclusion, in which Ebbinghaus (1913) writes, "Thus we have currently a great number of theories, which differ in their basic principles and also differ in the particular development of those principles." (p.54). This is very similar to the situation today. Most illusions produce multiple classes of explanation, and while we might hope that illusory phenomena would provide the ideal testing ground for different theories, we instead find that many theoretical approaches seem adaptable enough to prevent exclusion. For instance, the Müller-Lyer illusion spawned 12 theories between 1889 and 1902 (see Boring, 1942, chapter 7, for the list), and most of these theories could almost certainly find parallels in the current literature. More than that, perhaps the most common account of the Müller-Lyer illusion found in psychology and perception textbooks is the approach credited to Thiéry (1895) and popularized by Gregory (2009), that the differences in perceived line length result from the visual system's misapplication of perspective cues. The longevity of this account is surprising because multiple counterexamples have been produced that cannot be explained in this way (see the chapters by DeLucia [ch. 15] and Deregowski [ch. 2] in this volume). Nevertheless, the most common account does seem to fit a story consistent with other visual facts and therefore is very hard to completely disregard.

Indeed, the extent to which the field has been unable to eliminate longstanding alternative hypotheses is somewhat dispiriting. In 1898, in a review of the then-current state of the illusion research, Peirce attempted to find comprehensive unity among the seemingly scattered and unrelated facts by offering three accounts; the review ends with the optimistic question, "Which of these explanations shall we accept?" The three hypotheses he proposes all have corresponding accounts today: (a) a repulsion account (i.e., an interaction in which local lines are mutually antagonistic in the immediate vicinity of the point of intersection); (b) a perspective account motivated by Thiéry (1895); and (c) an account in which directional shifts result from eye movements. Needless to say, more than a century later, the field is still unable to reach a consensus answer to Peirce's question. Weintraub (1993) chimes the bells for pessimism:

> As reader, researcher, writer, and reviewer of manuscripts on these issues, I conclude that new findings never seem to resolve any theoretical issues. In words of few syllables—we aren't getting anywhere. . . . Ignoring the formulations of everyone else is what most of us do now, which is perhaps why the literature on illusions (a word I eschew) has always been in such a muddle. (pp. 236–237)

Thus a look at the literature continues to create the feeling expressed by Lindworsky (1929): illusions are enigmas felt as thorns in the flesh, presenting themselves "in broad daylight, as if to mock our inability to master them." (pp. 391–392)

However, if we take a step back, we can see that illusions, at the very least, show us the conditions where the visual system behaves unexpectedly. Illusions therefore say something interesting about how the brain processes information. It is hard to imagine how basic visual neuroscience discoveries like three classes of photoreceptors, lateral inhibition, center-surround organization, orientation columns, modularity of neural functions, and so on would have been treated if we did not have phenomena that seem, sometimes in retrospect, to foretell the physiological results. More than that, advances in imaging and electrophysiology are allowing us to investigate visual phenomena with new forms of analysis. The technology is creating new and exciting ways to consider how the brain constructs our perceptual world.

Many of the chapters in this book result from these ways of thinking about the brain, and many other chapters will help guide future research in these directions. Illusions compel humans to view the world with a sense of amazement and give us a better understanding about how the brain constructs a version of reality. We therefore conclude that Purkinje's (1819) optimistic dictum, "Visual illusions reveal visual truths" (cited by Wade & Brožek, 2001), still holds some sway. We hope that the chapters in this book will contribute to the fulfillment of Purkinje's vision and will further inspire the desire to explore, create, share, and explain.

REFERENCES

Boring, E. G. (1942). *Sensation and perception in the history of experimental psychology*. New York: Appleton-Century-Crofts.

Borges, J. L. (1964). *Other inquisitions (1937–1952)*. Austin: University of Texas Press.

Coren, S., & Girgus, J. S. (1978). *Seeing is deceiving: The psychology of visual illusions*. Oxford: Lawrence Erlbaum.

Da Pos, O., & Zambianchi, E. (1996). *Visual illusions and effects*. Milano: Guerini.

Ebbinghaus, H. (1913). *Grundzüge der Psychologie*. Leipzig: Veith & Comp.

Fineman, M. (1966). *The Nature of visual illusion*. New York: Dover.

Flynn, O., & Shapiro, A. G. (2014). A note concerning the relationship between the Adelson's Argyle illusion and Cornsweet edges. *Psihologija*, *47*(3), 353–358.

Gregory, R. L. (2009). *Seeing through illusions*. Oxford: Oxford University Press.

Goto, T., & Tanaka, H. (2005). *Handbook of the science of illusion*. Tokyo: University of Tokyo Press.

Lindworsky, J. (1929). Zum problem der Gestalttäuschungen. *Archiv für die gesamte Psychologie*, *71*, 391–408.

Luckiesh, M. (1922). *Visual illusions: Their causes, characteristics and applications*. New York: van Nostrand.

Newton, I. (1730). *Opticks or, a Treatise of the Reflections, Refractions, Inflections,and Colours of Light*. London: William Innys

Ninio, J. (2001). *The Science of illusions*. Ithaca, NY: Cornell University Press.

Oppel, J. J. (1854–1855). Über geometrisch-optische Täuschungen. *Jahresbericht des physikalischen Vereins zu Frankfurt am Main*, 37–47.

Pierce, A. H. (1898). Geometrical optical illusions. *Science*, *8*(206), 814–829.

Robinson, J. O. (1972). *The psychology of visual illusion*. London: Hutchinson.

Sanford, E. C. (1898). *A course in experimental psychology: Sensation and perception*. Boston: Heath.

Thiéry, A. (1895). Ueber geometrisch-optische Täuschungen. *Philosophische Studien*, *11*, 307–370.

Todorović, D. (2014). *Realnost i iluzije: uticaj konteksta u vizuelnom opažanju. Reality and illusions: Effects of context in visual perception* (in Serbian). Belgrade, Serbia: Filozofski fakultet, Univerzitet u Beogradu.

Tolansky, S. (1964). *Optical illusions*. Oxford: Pergamon Press.

Turner, R. S. (2014). *In the eye's mind: Vision and the Helmholtz-Hering controversy*. Princeton Legacy Library Series. Princeton, NJ: Princeton University Press. (Original work published 1994)

Vicario, G. B. (2011). *Illusioni ottico-geometriche: Una rassegna di problemi*. Venezia, Italy: Istituto Veneto di Scienze, Lettere ed Arti.

Wade, N. J. (1982). *The art and science of visual illusions*. London: Routledge & Kegan Paul.

Wade, N. J. (2006). *Perception and illusion: Historical perspectives*. Berlin: Springer.

Wade, N. J., & Brožek, J. (2001). *Purkinje's vision: The dawning of neuroscience*. Mahwah, NJ: Lawrence Erlbaum.

Weintraub, D. J. (1993). The logic of misperceived distance (or location) theories of the Poggendorff illusion. *Perception & Psychophysics*, *53*(2), 231–238.

Wundt, W. M. (1898) Die geometrisch-optischen Täuschungen. *Ahandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften, mathematisch-physische Klasse*, *42*(2), 55–178

Zusne, L. (1968). Optical illusions: Output of publications. *Perceptual and Motor Skills*, *27*(1), 175–177.

Zusne, L. (1986). Geometric-optical illusions: An update on the output of publications. *Perceptual and Motor Skills*, *62*(3), 729–730.

PART I
Introductory General Chapters

Chapter 1
Early History of Illusions

Nicholas J. Wade

INTRODUCTION

We tend to think of visual illusions in terms of the history of psychology rather than the history of vision. That is, we associate them with the illusion figures that were devised from the latter half of the 19th century onward. The alternative view, which is the one I adopt, is to consider them in the much longer history of vision; illusions were not as rigidly defined as in the 19th century, but they were widely discussed (see Wade, 2005a). Nonetheless, there are some commonalities in the approaches throughout these historical periods. Illusions of vision have proved more intriguing than the veridical perception of object properties. Since most of perception is veridical, it is the occasional departures from veridicality that have provided fascination. Illusions, or mismatches between perception of stimuli and their physical attributes, have been gauged by many means. The most venerable method has involved the comparison of percepts: when object properties appear to differ under different circumstances, then an illusion is said to have occurred. The only assumption that needs to be made is that the objects have not themselves changed between the two events. Physical benchmarks were introduced with an increased understanding of the physical world. Thus, perplexing percepts like the apparent bending of sticks when immersed in water could be given a more mundane interpretation when the laws of refraction were specified. However, the reference that has proved most attractive to many students of the senses is comparison between percept and underlying physiology. Visual illusions were studied long before geometrical optical illusions were given this name. For much of its long history, the study of vision has been confined to naturalistic observation. With the move from the natural world to the laboratory, the oddities of visual perception multiplied, and they received ever more detailed scrutiny.

The emphasis on errors of perception encompasses virtually all perceptual research from the ancients until the present. The constancies of perception have been taken for granted whereas departures from constancies (errors or illusions) have fostered fascination. The methods of recording observations have become more refined, but this has not resulted in an increased concern with veridicality. Rather, the range of illusions that are studied has exploded. Any study of illusions is predicated on an assumption of a standard from which the errors can be assessed. The standards themselves have changed over the centuries, largely as a consequence of developments in the physical and life sciences. Accordingly, the nature of perceptual error itself is examined here before surveying particular visual illusions.

Recording the consequences of perception started long before written reports were kept, and so art could be encompassed within this purview. Writers have remarked on their own visual experiences since the invention of script, so that a large body of descriptions of observations has accumulated. This body is dissected in the present chapter. Descriptions of visual experience are likely to be anchored in a more solid environment than the theories proposed to account for them because the theories themselves have been dependent upon concepts derived from other disciplines. The function of vision is to guide behavior, and insofar as this guidance is successful, there might seem to be little in the subject to warrant enquiry. Indeed, the eternally entertained theory of naive realism speaks to this issue—the world is as it is perceived. Nonetheless, there were circumstances in which the phenomena of vision were remarkable and remarked upon. Visual experiences in darkness (as a consequence of pressure or a blow applied to the eyeball) were not only remarked upon around 500 BC, but they could have provided the phenomenal source for emission theories of vision—that light issues from the eye itself. Such theories might seem fanciful to us now, but the phenomena upon which they are based are as readily experienced today as they were two and a half thousand years ago, and descriptions of them have been repeatedly refined throughout that period. Afterimages provide a similar example; they can be seen following exposure to bright light, they take on the shape of the intense stimulus, and they linger for an appreciable time. They could have acted, together with the reflections seen in water and in the eye, as a basis for the belief that vision was mediated by images or copies of external objects. These phenomena and many others require an interpretation by any adequate theory of vision; the theories might be supplanted, but the phenomena remain.

Science involves recording and interpreting natural phenomena. In contemporary science, the records are the results of experiments, and the many and varied phenomena

are posited in well-defined compartments, like physics, physiology, and psychology. These compartments are a relatively recent convention, as are the specialists who labor under titles such as physicists, physiologists, and psychologists. Neither the phenomena nor the practitioners were so clearly defined in the distant past. In antiquity science, if such it should be called, was based upon describing and classifying observations of naturally occurring events. The sense of science was sight, and vision itself was an integral part of the development of science. What was seen could be described, classified, and even subjected to mathematical analysis.

For most of us, our first contact with illusions is through magic, and it certainly reaches a larger section of the population than the illusions studied in science. We now accept that magic involves tricks that comply with natural laws, but this was not always so. Magicians of the past appreciated the laws of light and could manipulate attention with far greater subtlety than was the case for students of the senses (Brewster, 1832). They were able to beguile and bamboozle those who did not share their knowledge, which they kept as secret as possible. Magicians are applied scientists, using knowledge of illusions (and selective attention) with consummate skill without wishing to delve too deeply into their underlying bases. There is now renewed interest in the magician's art but from the standpoint of visual and cognitive science (Kuhn, Amlani, & Rensink, 2008; Kuhn & Tatler, 2011; Martinez-Conde & Macknik, 2008). This chapter reflects various strands of historical research on what we now consider to be visual illusions (Fig. I.1-1). However, despite the color illusion incorporated in the figure, consideration is restricted to spatial illusions.

In antiquity, Ptolemy (c. 100–170) and Ibn al-Haytham (c. 965–1039; also known as Alhazen) were concerned with general features of perception as well as illusions (Sabra, 1989; Smith, 1996). Ptolemy drew a distinction between subjective and objective aspects of visual phenomena and devoted considerable space to errors of perception. Indeed, he was one of the first writers to provide a detailed account of illusions; they are classified and then considered under the headings of color, position, size, shape, and movement. Ibn al-Haytham adopted a similar analysis of the errors of direct vision, although he extended the range of phenomena for which they occur. He was more explicit in categorizing the three modes of vision in which illusions can occur. Illusions were to be understood in terms of the breakdown of the process of inference. Nonetheless, the categories Ibn al-Haytham gave for the errors of sight were fewer than the visible properties he listed. Errors of inference were confined to distance, position, illumination, size, opacity, transparency, duration, and condition of the eye.

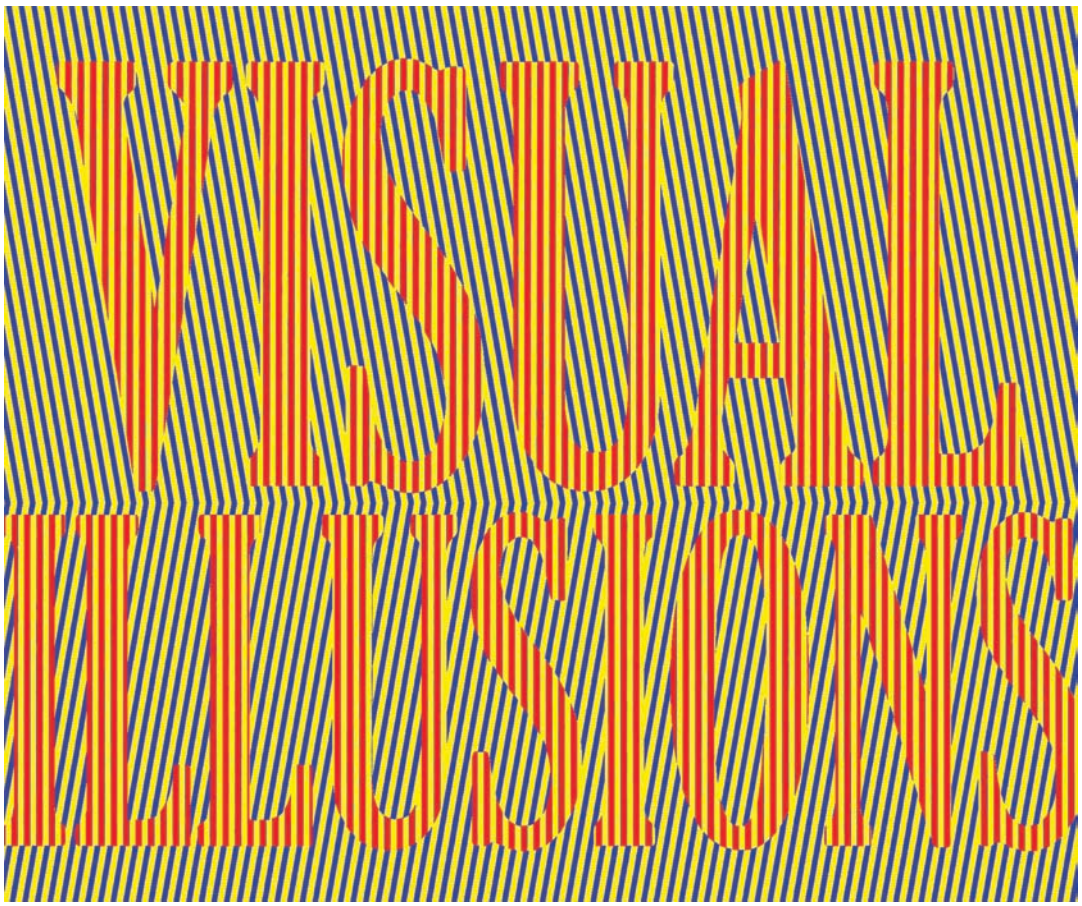


Figure I.1-1. *Illusions of Vision* by Nicholas Wade. All the lines defining the words are vertical but those for VISUAL appear to be tilted in the opposite direction to those for ILLUSIONS due to the background orientations. Color assimilation is also evident as the yellow background is physically equivalent throughout. (Image © Nicholas Wade.)

NATURE OF PERCEPTUAL ERROR

Illusions can be considered as errors in perception, and as such they were remarked upon before the basic perceptual processes were either described or appreciated. This was so because it was possible to compare observations of the same objects over time and to note any discrepancies between them. The Greek philosopher, Aristotle (c. 384–322 BC), did entertain the possibility of illusions entering into a particular sense. The examples he mentioned were those of color or sound confusions and errors in spatial localization of colors or sounds. The modern definition of "illusions" applies to differences between the perception of figures and their physical characteristics. Consensus concerning an external reality did not exist in antiquity, and so attention was directed to those instances in which changes in perception occurred—that is, when the same object appeared to have different properties under different conditions. According to this observational definition of illusions, all that is required is an assumption of object permanence; thereafter, changes in the appearance of the same object are classified as illusions.

An abiding example of this is the variation in apparent sizes of celestial bodies at different locations in their transits through the sky. These have been exemplified in the moon illusion—its larger appearance near the horizon than high in the sky (Fig. I.1-2). It is, of course, a size illusion, but it has also been interpreted as a distance illusion. The moon illusion presented an enigma in the past, and it is one that still persists (Ross & Plug, 2002). Modern attempts at explaining it remain problematical (Herschenson, 1989). It provides a quintessential example of illusion because the observations have been consistent but the interpretations have shown a progressive change: it was analyzed initially as a problem of physics, then physiology, and finally psychology.

The celestial illusion was known long before Aristotle. Plug and Ross (1989) describe a 7th century BC cuneiform inscription on a clay tablet from Nineveh that could be interpreted as describing the phenomenon. Aristotle's account was clear, and he related the variations in size to the effects of mist: "the sun and the stars seem bigger when rising and setting than on the meridian" (Ross, 1931, p. 373b). This was essentially the explanation given by Ptolemy in the *Almagest*, despite his appreciation that objects in a rarer medium appear smaller:

> It is true that their true sizes [of the celestial bodies] appear greater at the horizon; however, this is caused not by their shorter distance, but by the moist atmosphere surrounding the earth, which intervenes between them and our sight. It is just like the apparent enlargement of objects in water, which increases with the depth of immersion. (Ross & Ross, 1976, p. 378)

However Ptolemy wrote the *Almagest* before his *Optics*, in Book III, in which a psychological account of the phenomenon, in terms of apparent distance is given. There has been some dispute concerning the relationship between

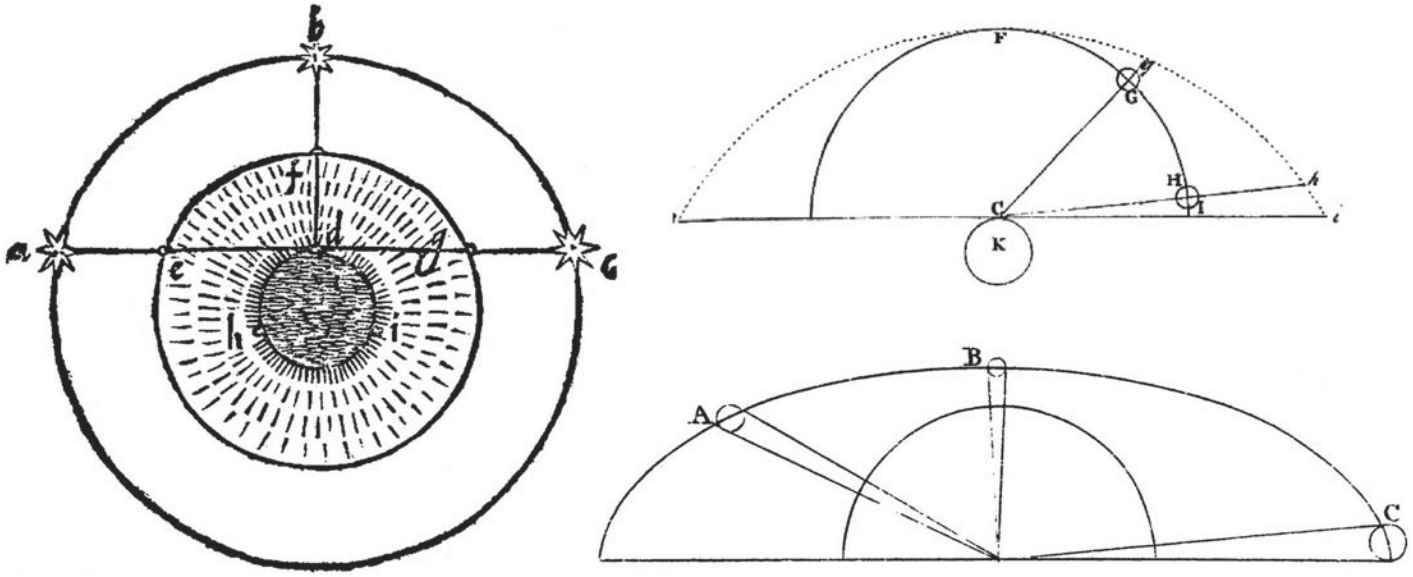


Figure I.1-2. Celestial illusions. Left: Galileo's (1606) diagram of the sun/moon illusion. Galileo wrote: "The sun and the moon appear larger by reason of that vaporous region when they are close to the horizon than when they are high in the sky, and hence you are convinced that their apparent enlargement depends upon that illumination. Both positions are true—the vaporous air is illuminated, and the sun and moon do appear larger when close to the horizon as a result of the vaporous region—but it is false that there is any connection between the two propositions or that the enlargement depends upon the illumination of that region" (1623, as translated in Drake & O'Malley, 1960, pp. 315–316). Upper right: A diagram of the moon illusion, produced by Desaguliers (1736) who wrote: "When we look at the Sky towards the Zenith, we imagine it to be much nearer to us, than when we look at it towards the Horizon; so that it does not appear Spherical, according to the vertical section EFGHI, but Elliptical, according to the section eFghi. . . . Now when the Moon is at G, we consider it is at g, not much farther than G; but when it is at H, we imagine it to be at h, almost as far again. Therefore, while it subtends the same Angle as it did before (nearly), we imagine it to be so much bigger as the Distance seems to us to be encreased" (pp. 390–391). Lower right: Young's (1807) diagram of the moon illusion: "The apparent figure of the heavens being nearly like the curve ABC, the sun or moon at A or C appears to be much larger than at B" (p. 787). (Image © Nicholas Wade.)

the two passages from Ptolemy (Ross & Ross, 1976; Sabra, 1987). Ibn al-Haytham absorbed both interpretations and acknowledged that both can apply but that differences in apparent distance were the principal cause of the illusion.

Galileo (1623) spoke at length about illusions of the senses, including celestial illusions, and his interpretation represented a break with Aristotelian psychology. Galileo's general interest in the senses was of a psychological and philosophical character; it reflected the fallacies and limits of senses and the ways in which scientific knowledge of the world could be gathered from potentially deceptive sensory appearances. This approach was applied by Galileo to celestial illusions; he had visited this phenomenon previously, under the pseudonym of Alimberto Mauri (Galileo, 1606). Not only did he argue that the sun appeared larger near the horizon because it passed through more vapors, but he illustrated the situation too (Fig. I.1-2). This represents a rare instance of Galileo applying an optical interpretation to a visual phenomenon (see Piccolino & Wade, 2008, 2014; Wade, 2007b; Wade & Piccolino, 2009). Paradoxically, despite his development of the telescope, Galileo did not contribute to the progress of optical science in his day, like that expounded by Johannes Kepler (1571–1630; 1604) and Christoph Scheiner (1573–1650; 1619).

With the invention of more powerful optical instruments in the 17th century, astronomers provided detailed measurements of the angular dimensions of celestial bodies. René Descartes (1596–1650; 1637/1902) commented that they subtend equal angles throughout their transit, and Molyneux (1687) gave the angular dimensions of the full moon. The most popular explanation remained one based on variations in apparent distance. Some attributed the misjudgment of distance to the visibility of interposed objects near the horizon, whereas others assigned it to aerial perspective. Berkeley (1709) devoted a substantial section of his *New Theory of Vision* to the moon illusion, particularly to Wallis's (1687) account, which he took as representative of the fruitless and unsatisfactory received opinion.

Berkeley (1709) stated:

> Now, between the Eye, and the Moon, when situated in the *Horizon*, there lies a far greater Quantity of Atmosphere, than there does when the Moon is in the *Meridian*. Whence it comes to pass, that the Appearance of Horizontal Moon is fainter, and therefore . . . it shou'd be thought bigger in that Situation, than in the *Meridian*, or any other elevation above the *Horizon*. (LXVIII)

Berkeley was adopting an interpretation that had been voiced, but usually rejected, by others. That is, the moon appears of lower contrast when near the horizon and is seen as larger due to this blurred or confused appearance. The interpretation extends back to Aristotle, but the size/distance theory was then more widely accepted. In the case of the moon illusion, Berkeley did cite some other work, and again he attacked the theories of the writers of *Optics*. The received view was that visual size was modified by visual distance, so that the horizon moon appeared larger because it was apparently more distant. Since Berkeley argued that neither distance nor size were immediately available to vision, it was illogical to adopt this view. He did, however, concede that a variety of factors could influence the apparent size of the moon, like the posture of the eyes or head. Faintness featured prominently in Berkeley's analysis, because it enabled him to account for variations of the apparent magnitude of the horizon moon under different degrees of clarity: it appeared even larger when faint than when clear.

Many of the apparent distance protagonists made informal observations of the horizon moon viewed with and without interposed objects. Plug and Ross (1989) ruefully remarked that "almost invariably the effect found by each writer was in line with his explanation of the illusion" (p. 19)! Ibn al-Haytham's suggestion that the vault of the heavens was flat resulted in attempts to represent it in the 18th century. Jean Théophile Desaguliers (1683–1744; 1736) and Thomas Young (1773–1829; 1807/2002) both produced diagrams implying its shallow dome and the manner in which this could be used to account for the moon illusion (Fig. I.1-2). Desaguliers and Young drew the spherical dome within the elliptical one, while Robert Smith (1689–1768; 1738/2004) adopted the opposite strategy in order to show the estimated difference in the magnitude of the illusion. Desaguliers was stimulated to conduct an experiment on changes in apparent size with the distinctness of objects.

PICTORIAL REPRESENTATION

Illusion played a significant role in the pictorial representations of objects, particularly when the Greek ideal of imitating nature was returned to in the Renaissance. Linear perspective was demonstrated by the architect and painter Filippo Brunelleschi (1377–1446) and formalized by a contemporary architect and mathematician, Leon Battista Alberti (1404–1472). Basically it was derived from the optics of Euclid (c. 323–283 BC). He analyzed vision in terms of a cone emanating from the eye (Fig. I.1-3). Perspective is the application of Euclid's visual cone to a glass plane intersecting it, and this device is now known as Alberti's window (Fig. I.1-3). Thus, the principles of reducing a three-dimensional scene to a two-dimensional picture were formulated before the image forming properties of the eye had been described. The optics of antiquity met the art of the Renaissance in the context of linear perspective. The technique of perspective painting was rapidly adopted by artists from that time on, and many textbooks described its rules. There are some excellent treatments of the history of perspective, but that by Kemp (1990) is especially instructive because it deals with the emergence of linear perspective in the 15th century and its subsequent development. Greek artists also applied some form of perspective, as Edgerton (1975) argued elegantly.

Linear perspective involves specifying a station point, picture plane, ground plane, and vanishing point enabling depiction of a single image of a scene. Representations of solid objects on two-dimensional surfaces, following the rules of linear perspective, were assisted by using the *camera obscura*. A rudimentary form of the instrument was described by Ibn al-Haytham, and Leonardo da Vinci (1472–1519) likened the eye to a camera (Keele, 1955; Kemp, 1990; Strong, 1979). Despite these fertile suggestions, the *camera obscura* was used extensively by artists rather than scientists, because the optics and anatomy of the eye were only dimly discerned before the 17th century.

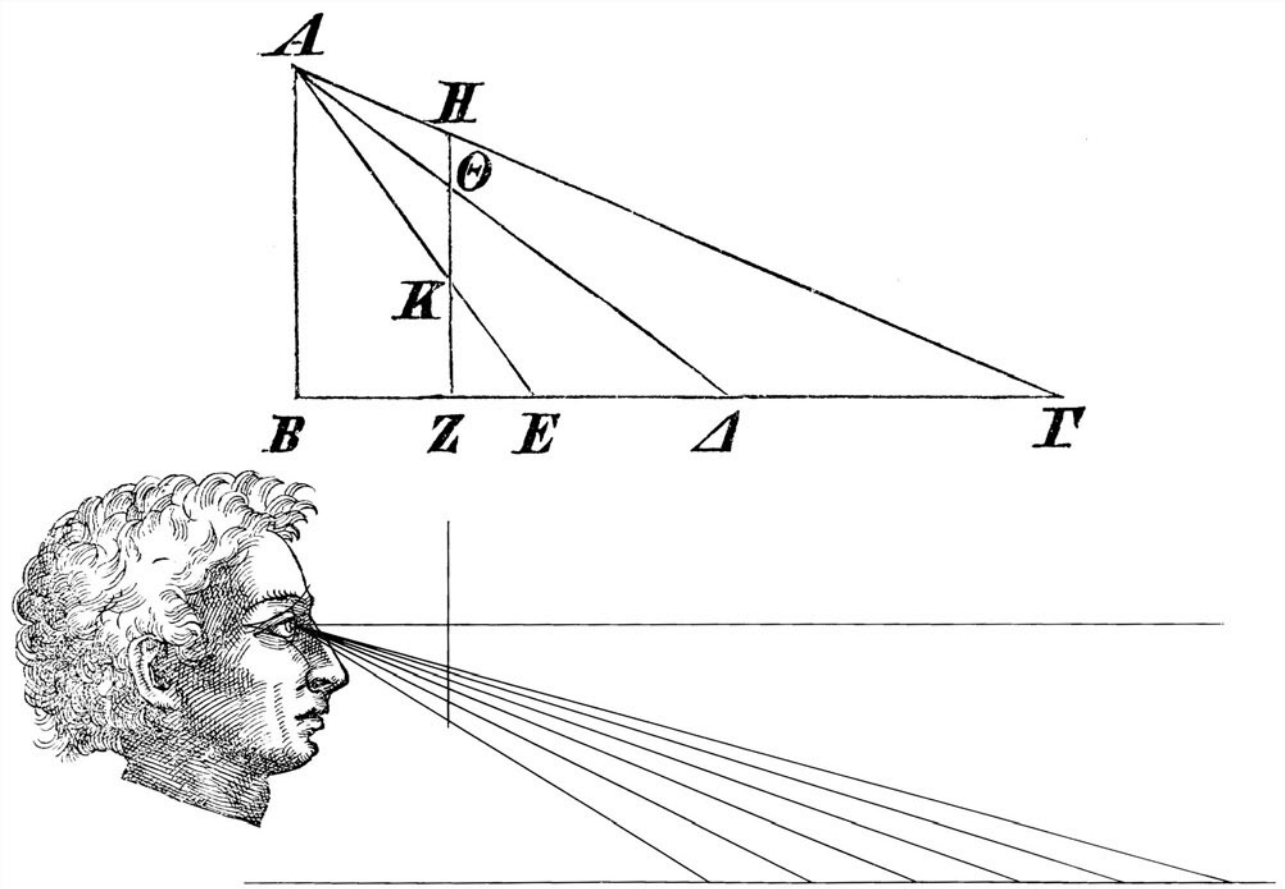


Figure I.1-3. Euclid's cone and Alberti's window. Upper: Euclid's diagram of projections from the eye A to points on a ground plane (from Euclid, 1895). Lower: A portrait of Leon Battista Alberti (after an engraving in Reusner, 1589) upon which is superimposed a drawing of rays from equally spaced tiles on the ground plane intersecting the vertical picture plane to meet the eye. (Image © Nicholas Wade.)

Leonardo saw not only the benefits that could derive from the application of the rules of linear perspective but also made the first systematic distortions of these rules. Soon after central perspective had been widely adopted in art, it was distorted in the form of accelerated and decelerated perspective architectures and anamorphic paintings. Descriptions of reversals of apparent depth were recorded long before the formalization of rules for linear perspective, although most attention was paid to reverse perspective in the 19th century; Wallin (1905) described these studies in detail. In anamorphic art the appropriate viewpoint differs from normal or perpendicular to the picture plane, so that the pictorial content can only be seen when the picture is viewed awry or through some appropriate optical device like a cylindrical mirror (Baltrusaitis, 1976). However, one of the most pervasive forms of manipulation has been the many and varied attempts to fool the eye (*trompe l'oeil*) with flat paintings. Successful examples of *trompe l'oeil* are rare, and those that do succeed usually place constraints on the viewer. For example, Samuel van Hoogstraten's (1627–1678) perspective cabinet in the National Gallery (London) has a single viewing aperture on either side, and Andrea Pozzo's (1642–1709) ceiling painting Apotheosis of St Ignatius in the church of St Ignatius (Rome) has a disc on the marble floor that defines the correct viewing position (Lingelbach, 2012b; Pirenne, 1970). These have been referred to as visual illusions rather than visual allusions to depth that are seen with conventional perspective pictorial images (Wade, 1990; Wade & Hughes, 1999). That is, they produce a unified and compelling percept of a depth that is not present on the painted surface; the conflicting cues provided by binocular vision and pictorial framing are either not operating (as in van Hoogstraten's cabinet) or are not applicable due to the dimensions and structure of the pictorial image (as with Pozzo's ceiling). After the Renaissance, linear perspective was one of the techniques of visual illusion that could be manipulated, as is evident from the remarks of Francis Bacon (1561–1626; 1625/1857). The mythical House of Salomon, described by Bacon in his *New Atlantis*, displayed all delusions and deceipts of sight as evidence of the advancement of science. Size and distance played a prominent role in such deceipts.

If there is an assumption of object permanence, then an illusion occurs when the same object appears to have different properties (of color, position, size, shape, motion, etc.) under different circumstances. Aristotle's description of the motion aftereffect (MAE) was based on a comparison of percepts: "when persons turn away from looking at objects in motion, e.g., rivers, and especially those which flow very rapidly, they find that the visual stimulations still present themselves, for the things really at rest are then seen moving" (Ross, 1931, p. 459b). The phenomenon was

presumably considered worthy of note because the stones at the side of the river appeared stationary prior to peering at the flowing water but not afterward (Wade & Verstraten, 1998). The changing perception of objects was the source of Aristotle's interest in phenomena like afterimages, aftereffects, color contrasts, and diplopia. It was precisely such variation in perception that led to the Platonic distrust of the senses.

Motion illusions, like the MAE and induced motion, were considered in detail before spatial illusions were investigated. The latter were a preoccupation from the late 19th century, when geometrical optical illusions took center stage in studies of vision. Moreover, the phenomena so called are often associated with a particular theoretical outlook. Empiricist philosophers, and those students of perception who followed them, set out from an ambiguous starting point. They assumed that the retinal image was static and impoverished and that something had to be added from past experience to remove the equivocality. Illusions could intrude during this amplification of the retinal information (Gregory, 2003, 2009). An illusion requires a yardstick or reference relative to which it can be assessed. In fact, underlying virtually all illusions are the mismatches of observation described earlier. However, in terms of interpretations the interest is directed to the source of the mismatch. The classical example is the apparent bending of a stick when immersed in water. There is a compelling contrast between the appearance of the stick in air and partially immersed, but does the stick change its characteristics? This provided a challenge to the assumption of object permanence, and alternative explanations were sought.

AMBIGUOUS FIGURES

Perceptual ambiguity refers to the alternation over time between differing interpretations of a pattern. Ambiguities often involve fluctuations in apparent depth that occur when insufficient evidence for veridical depth is available, either in pictures or when viewing distant objects. The latter is the condition under which Ptolemy described ambiguity: distant sails of ships appeared to change between appearing concave and convex (Smith, 1996). In contemporary vision research, ambiguities are studied almost entirely in the province of pictures; they feed from the flat surface. Visual ambiguities offer alternative interpretations of the depth in drawings or alternative descriptions of objects defined by equivalent contours (see Fig. I.1-4). They occur more readily with monocular observation, and the alternatives appear successively rather than simultaneously; their perception is referred to as multistable. Ambiguities seldom arise when relatively near solid objects are viewed.

We associate ambiguous stimuli with relatively recent research in perception, but pictorial ambiguities have been produced in Italy for many centuries. Roman mosaics display not only the Gestalt laws of grouping but also subtle pictorial ambiguities (Wade, 2012a). Geometrical mosaics, like those that have been found throughout the Roman world, often manipulate the ambiguous depth in flat designs (Dunbabin, 2006). They abound in the borders of mosaics and often assume positions of central importance, like the design shown in Figure I.1-4.

Other descriptions of depth fluctuations echo the conditions observed by Ptolemy. Smith (1738/2004) described and illustrated the apparent reversals of direction that occur with motion of the sails of distant windmills (Fig. I.1-4). David Rittenhouse (1732–1796; 1786) noted that the craters of the moon, viewed through a telescope, could appear concave or convex, depending on the direction of the shadows. He also observed similar instabilities when viewing near objects using a reduction tube; the ridges in chocolate bars could reverse. Albrecht von Haller (1708–1777; 1786) observed reversals in wax seals when illuminated from different directions; he also described similar effects due to the directions of shadows cast by specimens observed under the microscope.

Pictorial ambiguities were made explicit by many 18th-century painters including William Hogarth (1697–1764) and Pierre Crussaire (1749–1800). For example, Crussaire illustrated a vase defined by two different profiles (Fig. I.1-4). This phenomenon resurfaced in the psychological domain with Edgar Rubin's (1886–1951; 1915, 1921) vase/faces motif (see Pind, 2014). Many ambiguities are dependent on the lost dimension of pictures—depth. It introduces uncertainly in our vision, and this uncertainty is evidenced in the fluctuations in apparent depth that ensue. One of the most common modern examples of perceptual reversal is that of a hollow mask, and this was described by David Brewster (1781–1868) in 1826. Brewster also observed conversions of relief with cameos and intaglios, which depended on the direction of the illuminating source. Brewster (1826) referred to this as an optical illusion. Louis Albert Necker (1786–1861; 1832) corresponded with Brewster, who was an editor of *The Philosophical Magazine*, and described a perceptual alternation in a drawing of a crystal (Fig. I.1-4).

The Roman mosaic manipulations were not restricted to figure–ground alternations but engaged in spatial ambiguities of apparent depth too. The mosaic shown in Figure I.1-4 utilizes the lozenge (diamond-shaped) motif: the lozenges radiate from a central point and vary in color. They are symmetrical, with angles of 45º and 135º, so that eight lozenges radiate from a point to complete a circle. The outer boundaries of the lozenge-squares are bordered by chevrons, giving the impression of cubes with patterned faces. However, the cubes are spatially ambiguous because each lozenge can be a face of two cubes. Similar designs have been found in the House of Dionysius, Cyprus (Kondoleon, 1995), and in the mosaics of Pompeii (Lingelbach, 2008). In the latter, the corners of large squares define the edges of lozenges, which in turn define the sides of smaller squares at 45º to the sides of the larger square. The smaller square can be seen as the face of two cubes. Lingelbach has also drawn attention to an even more complex 3rd-century mosaic in Portugal; it is made up from white, blue, and brown tesserae, and the lozenges and squares are arranged to produce multiple organizations so that the surface itself appears to be uneven. These systematic manipulations of simple geometrical shapes to represent solid forms are testament to the knowledge of perceptual organization in the designers of the mosaics.

As is evident from Figure I.1-4 (upper left), Roman mosaics were not restricted to geometrical designs along the borders, and other beautiful examples of figure–ground alternation have been found that are echoed in Rubin's

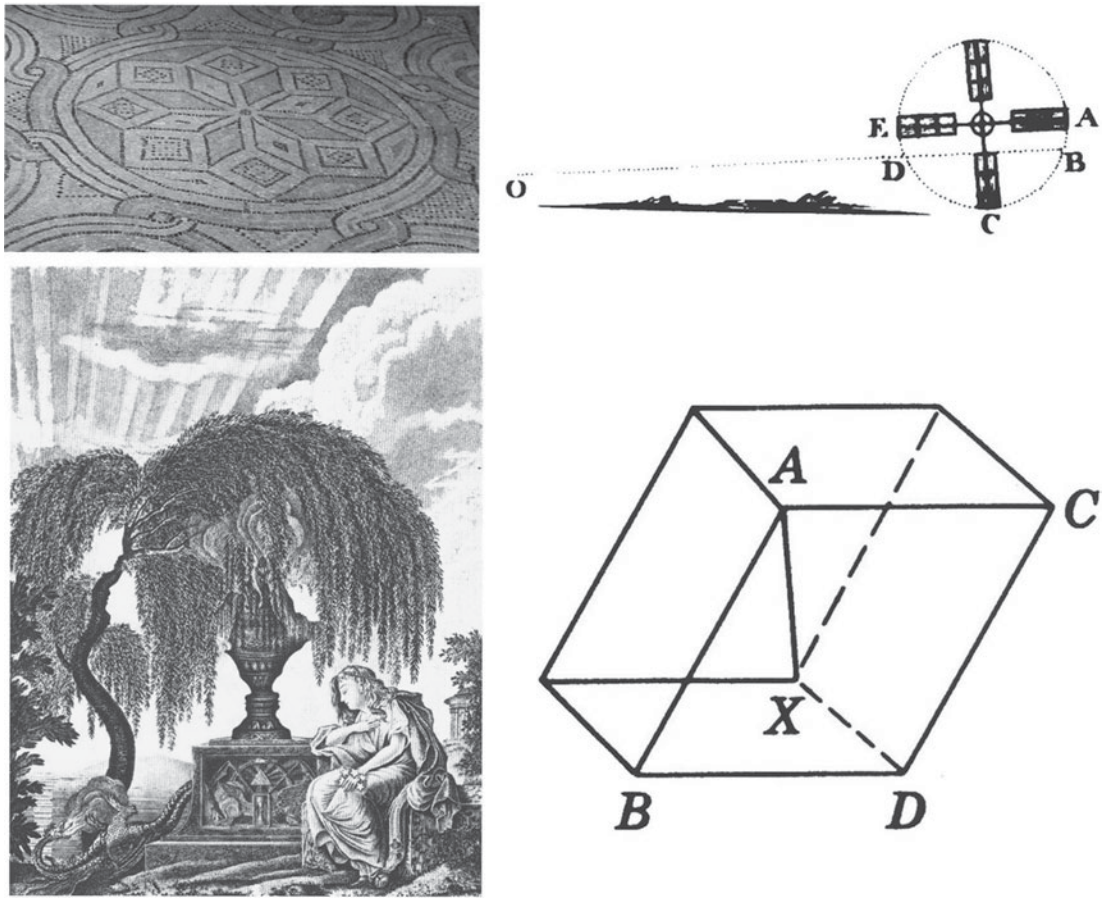


Figure I.1-4. Upper left: Photographic detail of a 5th- or 6th-century Roman mosaic in Domus dei Tappeti di Pietra, Ravenna, illustrating spatially ambiguous motifs; the eight radiating lozenges form sides of cubes, but a given lozenge could be the face of two cubes. The tesserae in the original are colored black, off-white, and light brown. Upper right: Smith's (1738) diagram of a distant windmill where the direction of rotation appears to reverse. Lower left: *The Mysterious Urn* by Pierre Crussaire (1795); the sides of the urn represent different profiles, and faces can be seen in the foliage too. Lower right: Necker's (1832) diagram of a reversing rhomboid. He wrote: "The rhomboid AX is drawn so that the solid angle A should be seen the nearest to the spectator, and the solid angle X the furthest from him, and that the face ACBD should be the foremost, while the face XDC is behind. But in looking repeatedly at the same figure, you will perceive that at times the apparent position of the rhomboid is so changed that the solid angle X will appear the nearest, and the solid angle A the furthest; and that the face ACDB will recede behind the face XDC, which will come forward; which effect gives the whole solid a quite contrary apparent inclination" (p. 336). (Image © Nicholas Wade.)

(1915, 1921) figures. He carried out most of his observations with simple patterns like Maltese crosses either in black and white or with arms defined by radiating lines and concentric circles (Fig. I.1-5, right). Rubin (1958) described the conditions under which alternation occurs in the context of figure and ground:

> when two fields have a common border, and one is seen as figure and the other as ground, the immediate perceptual experience is characterized by a shaping effect which emerges from the common border of the fields and which operates only on one field, or operates more strongly on one field than the other. (pp. 194–195)

In the diagram of a mosaic represented in Figure I.1-5, there are multiple ambiguities, with the patterns forming perceptual circles, squares, symmetrical arcs, and crosses. The manipulations of perceptual organization are far more subtle with such patterns than those given by the Gestalt psychologists in the 20th century (Wade, 2012a). Some mosaic floors are devoted solely to geometrical designs, often of astonishing complexity and variation.

Necker's name is often associated with reversing representations of cubes and Rubin's with a vase/face figure that alternates (Wade, Campbell, Ross, & Lingelbach, 2010). The possibilities of ambiguous interpretations of the same contours were both appreciated and manipulated by Roman mosaic artists as well as those of later periods (Piccolino & Wade, 2006a, 2007a). As noted earlier, Rubin (1915, 1921) did make a rather crude illustration of such a design (Fig. I.1-6), although more elegant examples proliferated more than a century earlier. In a footnote, Rubin remarked that he had seen earlier examples of what he called this joke, both in an actual vase and in a picture. Indeed, Rubin made many references to examples of figure–ground segregation in classical art and aesthetics, but he did not refer to the many examples of figural ambiguity that employed the structure with which his name is linked.

The use of vase outlines to define different profiles was a popular graphical theme in the late 18th century, and the ambiguity was often extended to the foliage of trees surrounding the vase or urn, as noted previously for Crussaire's mysterious urn (Fig. I.1-4, lower left). In fact,

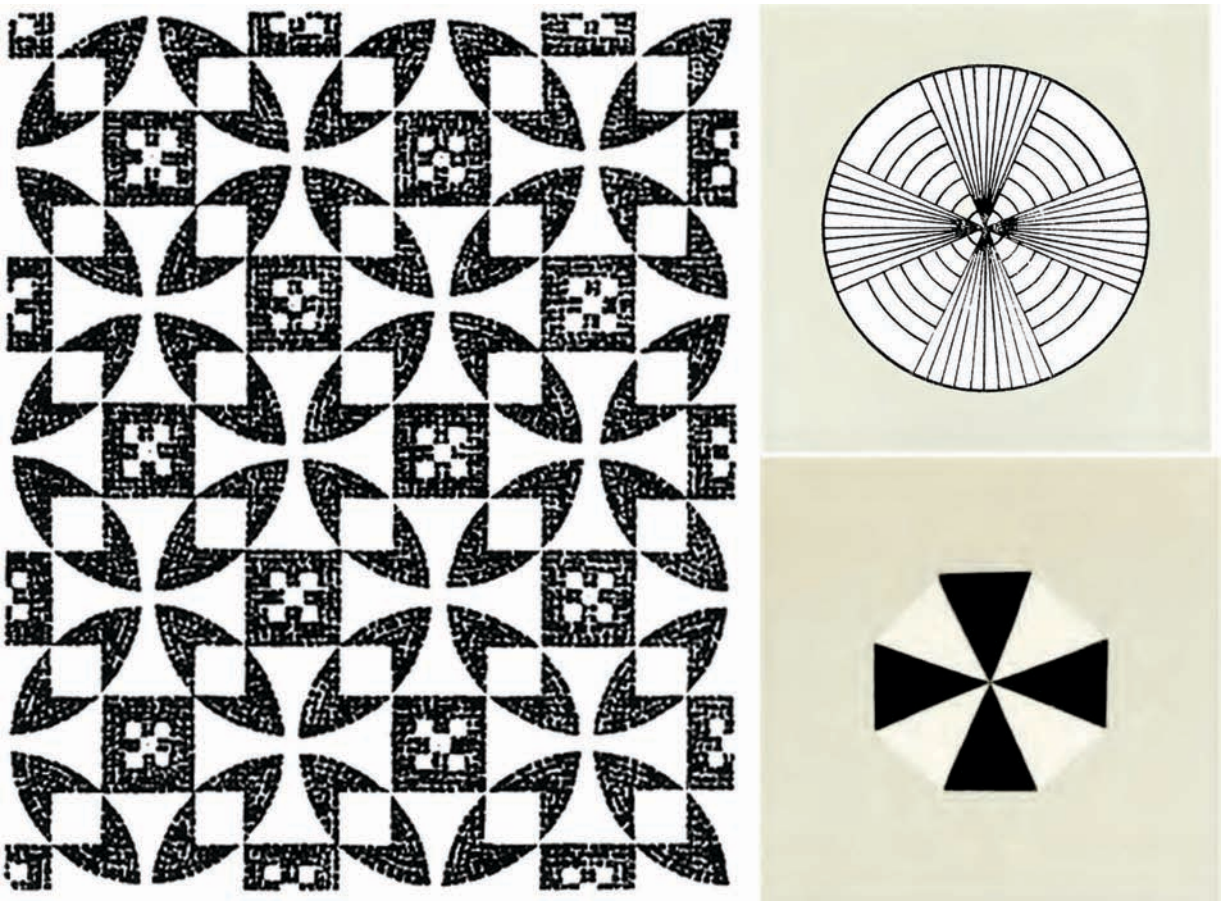


Figure I.1-5. Left: A geometrical figure/ground reversing design (based on a Roman mosaic) in which circles, squares, and crosses fluctuate in visibility (after an illustration in Campbell, 1988). Right: Rubin's (1921) simple cross figures used in his studies; they were black and white on a grey background. With regard to the upper right pattern he wrote: "One sometimes experiences the radially marked and sometimes the concentrically marked cross" (p. 38). The figure on the lower right was considered to be particularly good for producing alternations: "one can here see either an upright or an inclined cross" (pp. 53–54). (Image © Nicholas Wade.)



Figure I.1-6. Left: Rubin's (1921) vase/faces design. Right: Detail of a large banner hanging in the Chiesa dei Cavalieri di Santo Stefano; the facial ambiguity and various symbolic motifs can be seen at the top of the flag (from Piccolino & Wade, 2006b). (Image © Nicholas Wade.)

Rubin carried out almost all his experimental work with an alternating cross (Fig. I.1-5, lower right) rather than the vase/face motif with which his name is most closely associated (Fig. I.1-6, left); the latter was presented at the conclusion of his thesis almost as an afterthought. Crussaire's mysterious urn offers not only a vase/faces motif (with different faces on each side), but he also hides profiles in the foliage of the weeping willow. Moreover, the faces so concealed were politically potent, as they represented members of the deposed royal family at a time when displaying such allegiance was dangerous (Taws, 2007). Profiles and plants provided rich pickings for the artists who wished to amplify ambiguity or to conceal the identity of those so portrayed, and many examples were produced in the early 19th century (Dewitz & Nekes, 2002; Mannoni, Nekes, & Warner, 2004). Pictorial inversions were similarly employed to conceal recognition (Wade & Nekes, 2005).

Designs on flags and in other heraldic figures have not been accorded the recognition in the history of visual science that they warrant. The flag shown in Figure I.1-6 (right) hangs in the Chiesa dei Cavalieri di Santo Stefano, Pisa, and it provides an ambiguous example of ambiguity (i.e., the intention of the ambiguity is not known). The Ottoman flag was captured in a battle that took place in 1675 (Piccolino & Wade, 2006b), and the upper part suggests two facing profiles. The facial ambiguity is surprising because the unknown artist has flown in the face of the Islamic prohibition on representing human figures in religious decorations; this would certainly have applied to flags of the Islamic army and navy. However, the restrictions were not absolute, particularly in profane art, as long as the images did not represent God or the Prophet (Burckhardt, 1976).

When the flag is viewed upside down, the parts representing profiles can also take on different appearances. The central image, between the two facial profiles, could be identified as an inverted mosque peaked by the crescent moon or an Ottoman helmet. Thus, there could have been an element of orientational as well as figural ambiguity in the design. There are some 18th-century examples of the vase/face motif that are inverted, so that the facing faces are yet more difficult to detect. Contemporary accounts of perceptual ambiguities can be found in Alais and Blake (2005).

During the Renaissance, tiles replaced mosaics as the elements for floors, but similar geometrical themes were played, usually with differently colored tiles of marble. The churches and charterhouses of Italy are replete with tilings of intricate complexity. The examples shown in Figure I.1-7 can be seen in the Certosa di Calci, near Pisa; they date from the 17th century. Ambiguous representations of cubes could be constructed from four-sided tiles having the same shape, with opposite angles of 120° and 60°. An example is shown in the upper left illustration of Figure I.1-7: the

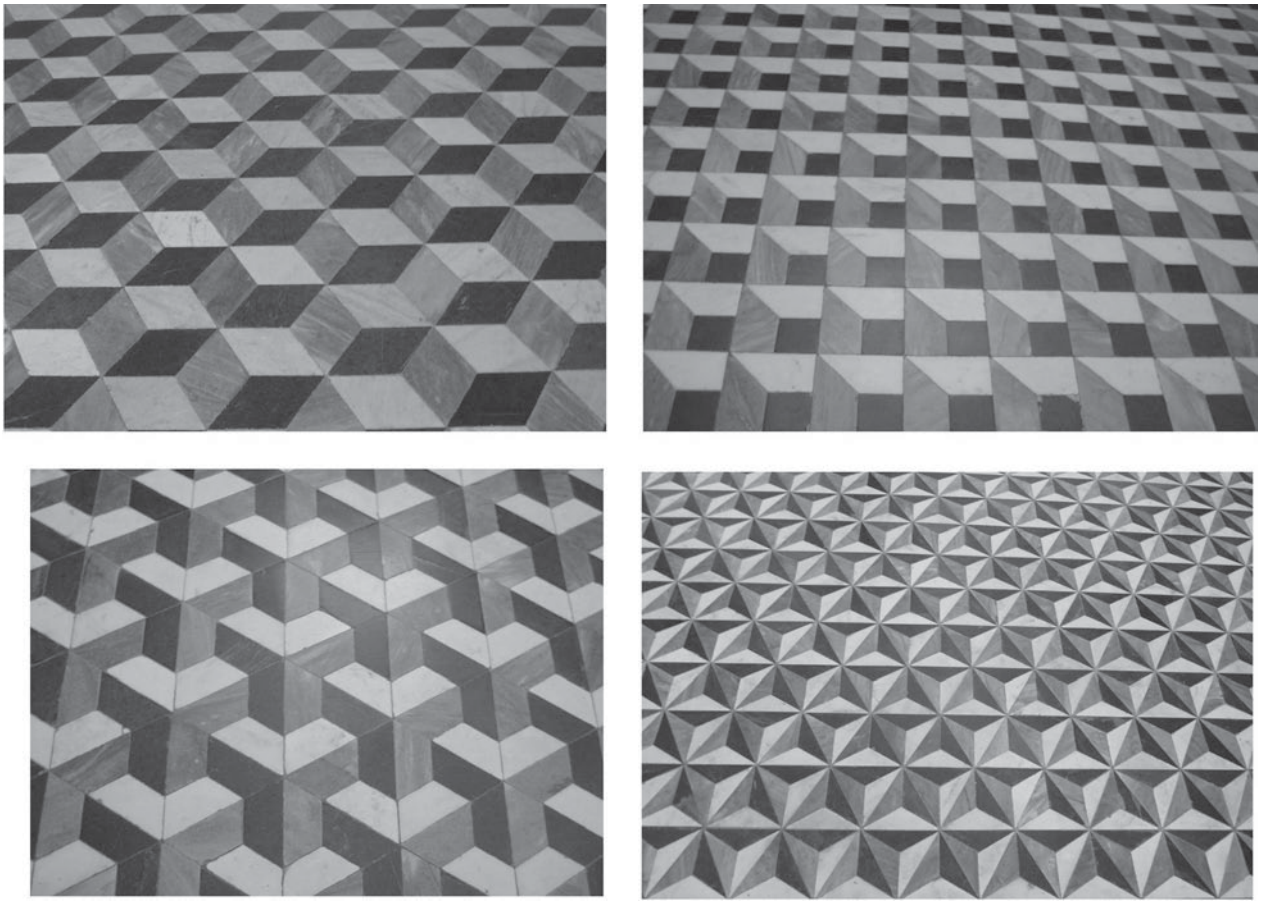


Figure I.1-7. Photographs of marble-tiled floors from Certosa di Calci, near Pisa, showing regular patterns of tiles distinguished either by the shape or shade of the marble. (Image © Nicholas Wade.)

three faces have different colors, and a given face can be part of two cubes with different perceptual orientations. The configuration is like that called the Beaunis cubes, and essentially the same patterns can be found in Roman mosaics (see Dunbabin, 2006; Lingelbach, 2008). The other examples in Figure I.1-7 are more complex both structurally and perceptually. The floor depicted in the upper right has two structural elements, squares and quadrilaterals with two corners at 90° and the others at 60° and 120°. It induces a strong impression of depth that is ambiguous: the dark square can appear either above or below the flanking quadrilaterals. The tiled floors shown in the lower figures are made up of elements having the same shape but differing in the shades of marble, so that they give the impression of solidity. That on the left has elements with quadrilaterals having angles of 60°, 90°, 90° and 120°, whereas the one on the left is comprised of triangles with internal angles of 30°, 30°, and 120°. There are many perceptual configurations that emerge when viewing it. The tiled floors found throughout Europe, especially in Italy but also in Great Britain, Turkey, Spain, Portugal, and Germany, are based on squares, rhombi, parallelograms, hexagons, or octagons, but none of the patterns utilize pentagons (Lingelbach & Wade, 2010). Despite the fact that the five-point star is one of the oldest known symbols, pentagons are very rarely used as decorations. There are some to be seen on the columns of Monreale Cathedral, Sicily (built in the late 12th century), which is known for its synthesis of Arabic, Byzantine, and Norman styles of architecture and art. The arches between the columns are made up of pentagons decorated with darts and kites. Thus the 12th-century artists seemed to know about the nonperiodic as well as periodic tiling.

Notwithstanding the long history of pictorial ambiguities before Necker, it was his description of it that stimulated further research. Charles Wheatstone (1802–1875) discussed the reversal of the Necker rhomboid in his classical memoir on the stereoscope (Wheatstone, 1838); he noted that the reversals were more frequent with monocular than with binocular vision. Indeed, the figure should be called the Wheatstone cube, and it was referred to in this way by Wallin (1905) in his detailed survey of reversing figures. Wheatstone was the first to use a skeleton three-dimensional cube rather than a drawing of a rhomboid and the first to note the differences in the perceived sizes of the near and far faces when reversals took place. Despite Wheatstone's observation of perceptual alternation in a solid structure, most of the ambiguous figures that followed were flat drawings, many of which show a striking resemblance to the mosaics and tiles described earlier. It was in the context of reversals of relief in cameos and intaglios that Schröder (1858) illustrated his steps or staircases (Fig. I.1-8). Ernst Mach (1838–1916; 1866) reduced the complexity of drawings like Schröder's staircase even further by presenting the figure that is frequently referred to as the Mach book (Fig. I.1-8): it looks like an open book with the spine either toward or away from the viewer. Other variants were produced and compiled in Jastrow's (1899, 1900) popular publications.

Mach (1866) did follow Wheatstone's lead and modified this simple arrangement to produce more dramatic fluctuations in apparent depth:

> I find it instructive to observe solid objects with one eye and to reverse my perception of them . . . If I fold a rectangular piece of stiff paper along a straight line to form two congruent rectangular parts and place it on a table in front of me so that when vertical it has the projecting side turned toward me, I obtain the image shown in Figure [8]. (Ratliff, 1965, p. 295)

Readers are encouraged to try this simple but potent demonstration for themselves; all that is required is some stiff white paper that can be folded as described. Mach was principally interested in the variations of perceived brightness with the depth fluctuations, but he did also comment on the changes in apparent shape:

> Furthermore, the form of the paper has changed. The lines *bf, bd* have become longer than *ae, ac*, the lines *cd* and *ef* shorter than *ab*. Finally, the corner *ab* no longer remains vertical, but the paper lies like an open book on a lectern before me. (Ratliff, 1965, p. 296)

When the depth of the card does reverse, any lateral head movements result in marked motions of the peaks and troughs.

Toward the end of the 19th and beginning of the 20th centuries, numerous ambiguous figures were produced in the context of art, science, and popular amusement; examples can be seen in Mannoni, Nekes, and Warner (2004). However, these tended to be figural rather than spatial. That is, the ambiguity involved difference of interpretation of the pictorial referents rather than their spatial locations. Three examples are shown in Figure I.1-8. Two were first produced in Germany, but they were popularized in America, and the third was printed in New York. The duck/rabbit ambiguous figure was published in a German weekly magazine (*Fliegender Blätter*) in October 1892 under the query "Which animals resemble one another most?" and above "Rabbit and duck." The illustration was copied in *Harper's Magazine* the following month, but it became more widely known after it was reproduced by Jastrow (1899, 1900) and is often attributed to him (Brugger, 1999). Jastrow did not change the ambiguous figure when he reproduced it, unlike Hill (1915), who simplified the young girl/mother-in-law picture. It was Hill's version that was popularized by Boring (1930). Wright (1992) suggested that Hill copied the figure from French sources, but an earlier version was published in Germany as a postcard in 1888. The somewhat satanic lower central illustration was printed in New York around 1900; the ambiguity is more complex as the two standing ladies in conversation with one another assume postures in which their profile faces and clothing define the ferocious frontal face.

Play upon the theme of the Mach card phenomenon has resulted in some intriguing artwork by Patrick Hughes; he calls them reversepectives (Papathomas, 2007; Wade & Hughes, 1999). Hughes has painted on protruding planes (like truncated pyramids and wedges) so that the parts that are physically close to the observer are pictorially distant. That is, the lines that would converge on a flat picture plane to allude to distant objects are physically closer in these works. They appear as flat paintings until the observer moves, whereupon they undergo a plastic motion that is beguiling. The pictorially distant (but physically

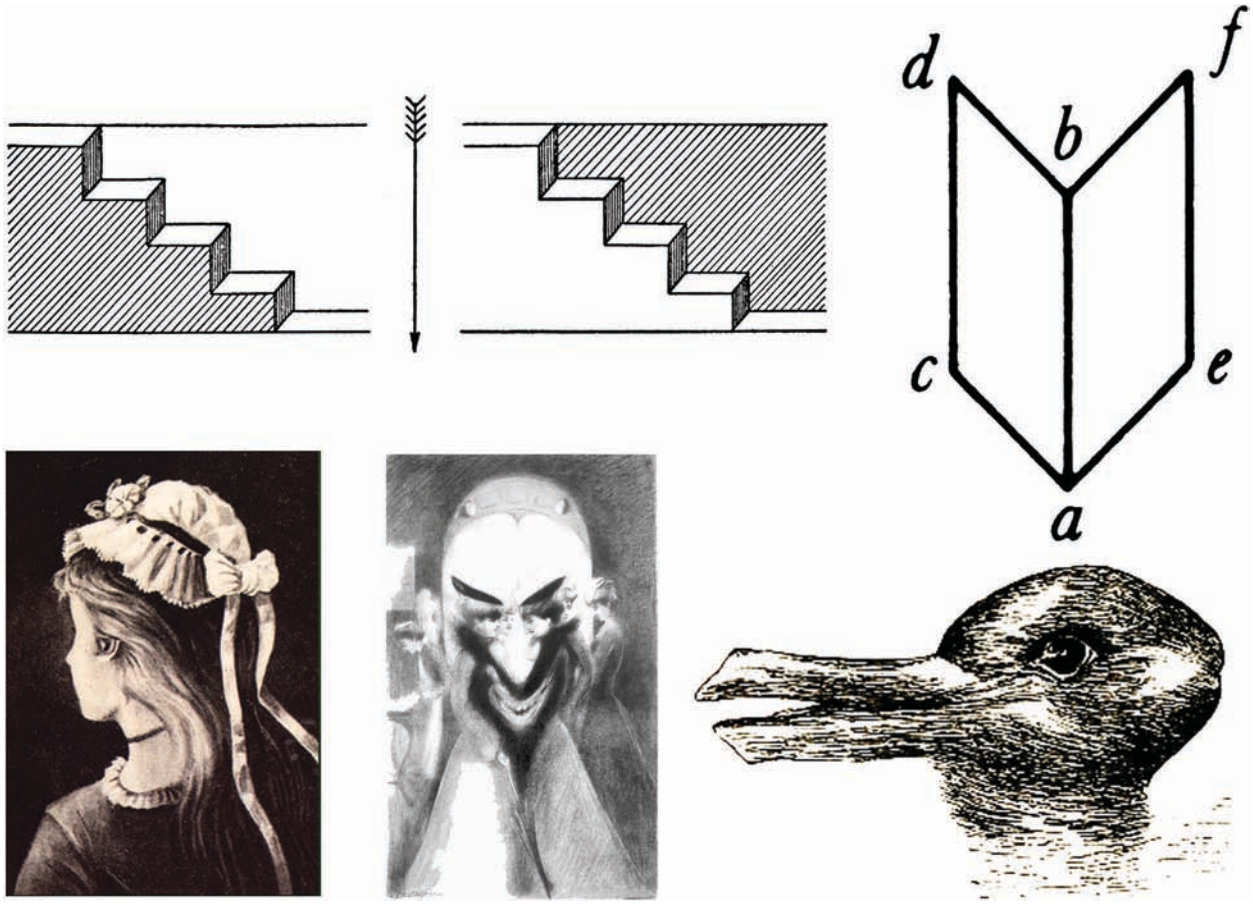


Figure I.1-8. Ambiguous figures. Upper: Schröder's (1858) staircase; the arrow indicates the direction of light so that the shaded areas would determine whether the stairs were seen from above or below. Mach's (1866) book illustration, that appears to fluctuate in depth but the changes are far more dramatic when card is constructed in three dimensions. Lower: The young girl/old woman figure as presented on a German postcard from 1888; it was modified by Hill (1915) and popularized by Boring (1930). *Gossip*, a print showing ladies talking with one another in postures such that their clothing defines a satanic visage; it was produced in New York around 1900 (reprinted with the kind permission of Werner Nekes). The duck/rabbit figure (from Jastrow, 1899) as it was printed in the *Fliegender Blätter* of 1892. (Image © Nicholas Wade.)

closer) parts move in the direction opposite head movement, whether this is left and right or upward and downward; the pictorially near (but physically more distant) parts move in the same direction as the head. A three-dimensional model of a Hughes work can be downloaded, printed, and constructed from Wade and Hughes (1999). Other artists have produced ambiguous three-dimensional structures so that they appear totally different from different vantage points (Fukuda, 2005; Raetz, 2002). For example, Fukuda has made a structure that appears like a violinist from one position and a pianist when viewed at right angles. The works by Hughes combine depth reversals, as with the spatial ambiguities mentioned earlier, with motion illusions, and it is to the latter that we now turn.

MOTION ILLUSIONS

The illusions examined here are stroboscopic motion, induced motion, MAEs, visual vertigo, and autokinetic effects. All but one of these phenomena were initially observed in the natural environment, and they have subsequently been studied under laboratory conditions. Stroboscopic motion is the exception. It is dependent upon visual persistence, which had long been observed, but it benefited most from the novel instruments: apparent motion was visible in the laboratory in a manner that had not been seen in the natural environment.

The initial interpretations of motion illusions were often in terms of eye movements. It reflected a desire to link an objective expression of motion (of the eyes) with a subjective experience of motion. One consequence of this endeavor was the more detailed understanding of how the eyes actually moved and the independence of most motion illusions of them. Eye movements are not correlated with induced motion, the MAE, or stroboscopic motion, but they do correspond to the direction and extent of visual vertigo. The extent of apparent visual motion in the autokinetic effect far exceeds that of involuntary eye movements, but there might be a weak relation between them. If eye movements are to provide a reasonable account of motion illusions, then they should apply to all the contours that are visible. This does apply to visual vertigo but not to induced motion, the MAE, and stroboscopic motion, all of which involve relative motion in a visible display. The autokinetic phenomenon involves the perception of isolated stimuli, and so interpretations in terms of some internal eye movement signals remain sustainable.

Visual Persistence and Stroboscopic Motion

Motion illusions involving visual persistence were among the first spatiotemporal phenomena subjected to quantification. The basic procedure was initially described by Ptolemy in the 2d century and in more detail by Ibn al-Haytham in the 11th century. A rapidly moving flame will be seen in positions it no longer occupies; if it is rotated rapidly then a circle will be seen. Leonardo da Vinci repeated this observation and demonstrated that a similar effect occurs when the eye moves with respect to a stationary flame (MacCurdy, 1938). Isaac Newton (1642–1727; 1717), in the second edition of his *Opticks*, used this phenomenon to estimate the duration of visual persistence; he suggested it was less than a second. More precise measurements were made by Chevalier Patrice D'Arcy (1725–1779; 1765): he built a machine with rotating arms onto which a glowing coal could be attached. By measuring the velocity required to complete a visible circle of light (in an otherwise dark room), he calculated the duration of visual persistence to be 8/60 s, or about 130 ms.

At the beginning of the 19th century, visual persistence was enlisted to produce a bewildering variety of visual motion illusions. The phenomenon was manipulated in three ways (a) by rendering visible rapidly moving lights during continuous viewing. (b) by successive stimulation of the eyes by slightly different figures, and (c) by briefly presenting parts of a single figure in rapid succession.

Wheatstone's (1827) phonic kaleidoscope or kaleidophone (Fig. I.1-9) rendered visible the paths of rapidly vibrating rods. The kaleidophone was an extension of a method described by Young (1800), in which silvered wire was attached to a piano string so that its vibration could be observed with the aid of a magnifying glass. Wheatstone, whose background was in musical instrument manufacture, constructed the kaleidophone to amplify the vibrations so that they could be seen with the naked eye. Silvered glass beads were attached to the ends of rods with different cross-sections and shapes; when the rods were bowed or struck, Chladni figures could be seen in the light paths traced by reflections from the beads. The kaleidophone was but one instrument of several that appeared in London during the 1820s. In 1825 John Ayrton Paris (1785–1856) described the thaumatrope, or wonder turner, which was even simpler in design (Fig. I.1-9), but he did not publish a description of it until 1827. A circular piece of card had different drawings on each side, and its ends were connected by string. When it was whirled, both designs were seen superimposed: birds could be seen caged during rotation but free when the disc was stationary; fragments of words written on each side of the disc could be rendered complete due to visibility of their persisting parts.

The impetus for these devices derived from observations made of the motions of spoked wheels behind or in front of railings. The initial description appeared in a brief note over the initials JM (1821; possibly the publisher John Murray): "When a spoked wheel, such as that of a carriage, or the fly of an engine, is viewed in motion, through a series of vertical bars, spokes assume the peculiar curvatures which are represented" (pp. 282–283). Peter Mark Roget (1779–1869), better known for his *Thesaurus* than for his experiments on vision, was fascinated by this phenomenon

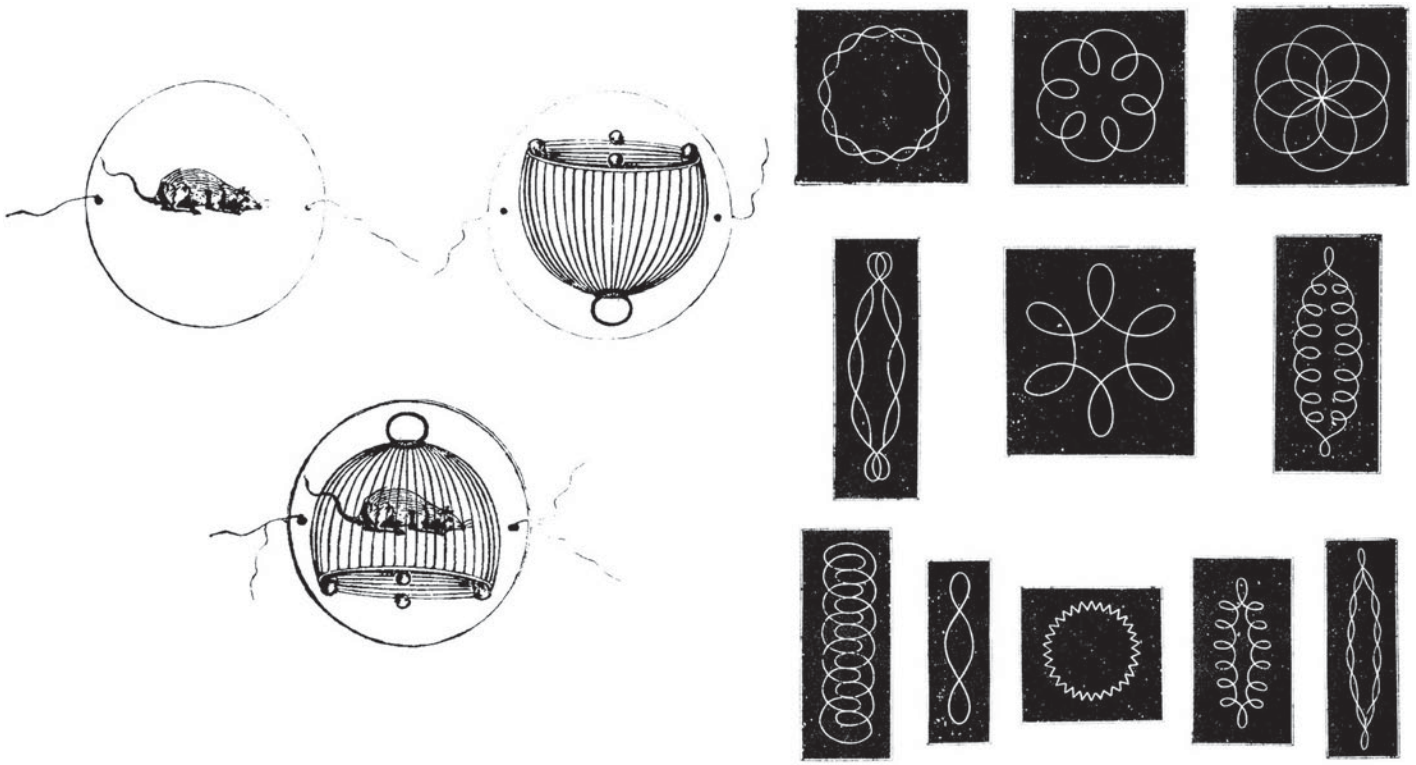


Figure I.1-9. Left: A thaumatrope or wonder-turner; the designs are printed on opposite sides of a disc so that when it is rapidly rotated they combine as shown below (derived from Paris, 1827). Right: Patterns of light produced by beads of glass on the rods of the kaleidophone (Wheatstone, 1827). (Image © Nicholas Wade.)

(Wade, 2010a). In 1825 he provided illustrations and a mathematical analysis of the phenomenon, relating it to persisting visual images. In the conclusion to his article he observed that "it might therefore, if accurately estimated, furnish new modes of measuring the duration of the impressions of light on the retina" (Roget, 1825, p. 140).

London scientific society was intrigued by the phenomena as well as by the instruments, and the fashion ensnared many whose names are not normally associated with toys. For example, Michael Faraday (1791–1867; 1831) cast his scientific eye over the effects and wrote a very influential article on optical deceptions. He was disparaging about Paris's thaumatrope, referring to it as a schoolboy trick, but he was attracted by Roget's analysis of rotating spokes and by his own observation of counter-rotating cogwheels. He constructed a simple arrangement of cut-out sectored discs to examine the effects further. Faraday (1831) wrote:

> The eye has the power, as is well known, of retaining visual impressions for a sensible period of time; and in this way, recurring actions, made sufficiently near to each other, are perceptibly connected, and made to appear as a continuous impression. (p. 210)

This statement excited the interests of others to construct instruments that could synthesize motion from a sequence of discrete images. In 1833 both Joseph Plateau (1801-1883; 1833), with his phenakistoscope or fantascope, and Simon Stampfer (1792–1864; 1833), with his stroboscopic disc, developed similar instruments for presenting a series of slightly different still pictures in rapid succession (Fig. I.1-10). Plateau (1833) described the instrument as

> a cardboard disc pierced along its circumference with a certain number of small openings and carrying painted figures on one of its sides. When the disc is rotated about its centre facing a mirror, and looking with one eye opposite the openings . . . the figures are animated and execute movements. (p. 305)

Stampfer's stroboscopic disc was very similar to Plateau's phenakistoscope, and both acknowledged the stimulus provided by Faraday's article. The issue of priority of invention inevitably ensued, and it is generally accorded to Plateau (Boring, 1942). However, Roget (1834) suggested in his *Bridgewater Treatise* that he had made such a device even earlier (Wade, 2011, 2012b).

Plateau (1833) appreciated that there were limits to the visibility of such apparent motion: if the rotation was too slow then each individual figure was seen; if it was too fast then they were all seen together in a confusing conglomeration. These instruments could be used by just one person at a time, whereas William Horner (1789–1837; 1834) developed a variant for group viewing: it consisted of a cylinder mounted on a vertical axis, with slits at regular intervals and a sequence of drawings on the opposite inside surface of the cylinder. He called it the dædaleum, but it became widely used in the latter half of the 19th century under the name of zoetrope (Carpenter, 1868).

Stroboscopic discs presented stimuli discretely, briefly, and in succession; that is, a sequence of drawings differing slightly from one another were viewed successively through slits in a rotating disc. To the astonishment of observers, a single figure appeared in motion: perceived movement was synthesised from a sequence of still pictures. Stroboscopic discs were used to study visual persistence and apparent motion, and Jan Evangelista Purkinje (1787–1869) made a variant of one in 1840; he called it the phorolyt or kinesiskop, and it was sold commercially as a magic disc (Matousek, 1961). Purkinje used his phorolyt to produce dynamic images of a range of natural movements generated from a sequence of static drawings. These varied from the pumping action of the heart to the walking movements of newts; he also used it to display photographs of his own rotating posture (Wade & Brožek, 2001).

Successive exposure of a moving window in front of a fixed scene has received sporadic description since the time of Leonardo da Vinci. A more detailed description was given by Robert Darwin (1786)—the father of Charles Darwin. He made a simple paper sail that he could rotate. Wheatstone's (1827) examination of visual persistence with the kaleidophone had alerted him to the possibilities of such a device, and he constructed one to demonstrate this effect more precisely: a sector in a metal disc exposed part of a design painted on glass in transparent colors; when the sector was rotated rapidly enough, the whole design was visible.

In the visual instruments described here, an aperture moved in front of a scene or picture. In 1829 Plateau investigated the effects of moving a pattern behind both moving and fixed apertures (Plateau, 1829, 1836, 1878); the instrument was similar to the phenakistoscope, involving rotating discs, and he gave it the name anorthoscope. Regular geometrical patterns moving behind a fixed aperture appeared deformed, or deformed figures appeared regular, like anamorphoses. The device was elaborated by Johann Karl Friedrich Zöllner (1834–1882; 1862), who demonstrated the importance of pursuit eye movements to the perceptual distortions. The aperture was typically a vertical slit, and the shapes were moved horizontally behind it. Simple shapes, like circles, appeared like ellipses with the long axis vertical for fast movements and horizontal for slow.

Induced Motion

Induced motion refers to the apparent visual movement of a stationary object, usually in the opposite direction to the real movement of other objects. It can be produced with a wide range of stimulus configurations, and it occurs with linear and rotary motions in the frontal plane and with motion in depth. Induced motion is a contrast effect in the motion domain: if several objects move together and another remains stationary, then the latter will appear to move in the opposite direction. What seems to be a clear description of induced motion was given over 2,000 years ago by Euclid (c. 323–283 BC) in his *Optics*. He wrote: "*If, when certain objects are moved, one is obviously not moved, the object that is not moved will seem to move backward*" (Burton, 1945, p. 371; italics in original). This account is rather puzzling, since it flies in the face of Euclid's geometrical account of vision (Wade, 1998). That is, Euclid sought to provide a theory of space and motion perception that reduced these to projective geometry; size perception was equated with visual angles and motion perception was associated with

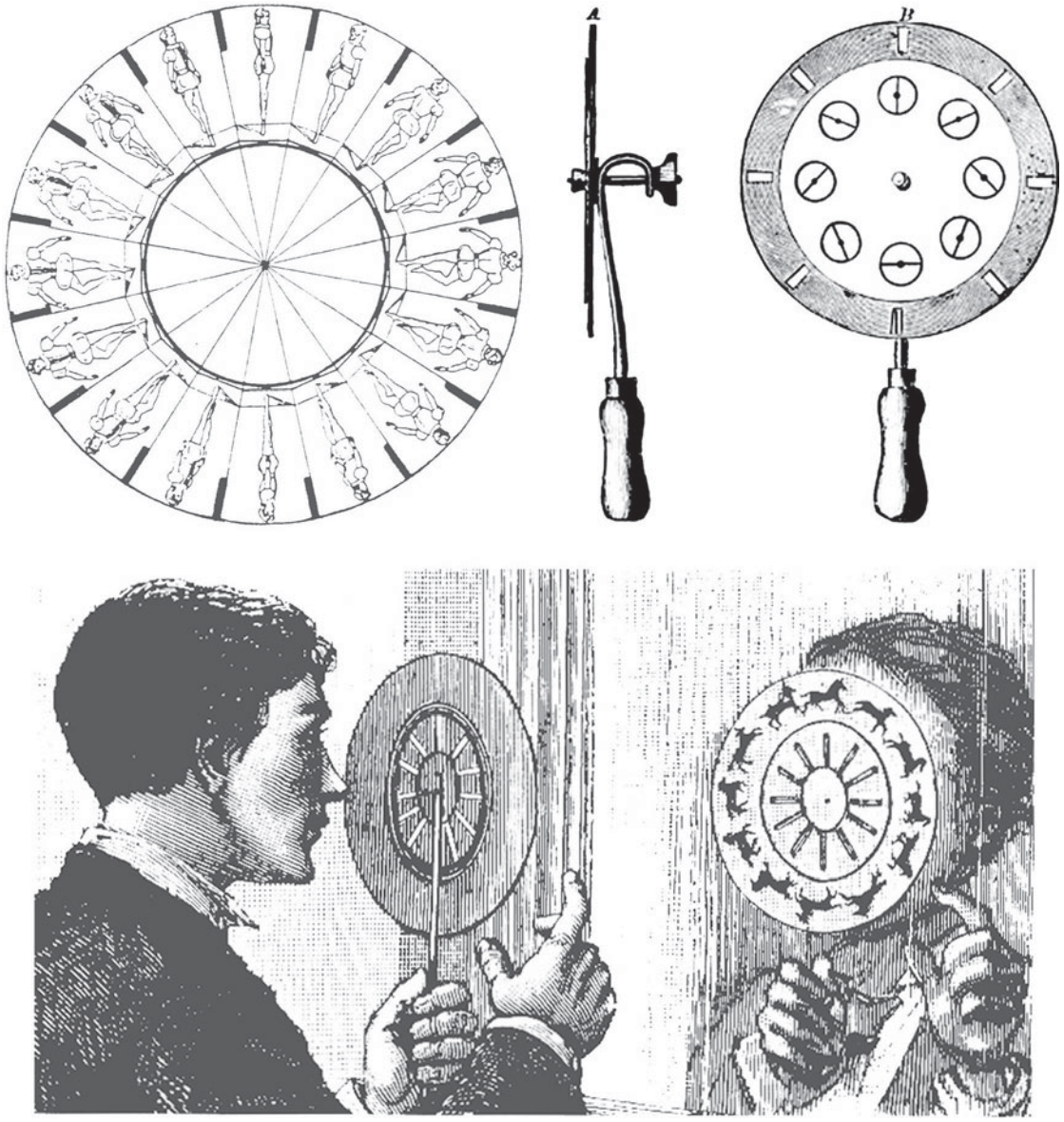


Figure I.1-10. Plateau's pirouetting dancers that could appear to move in the phenakistoscope or stroboscopic disc when rotated and viewed in reflection. The dancers are from Plateau (1833), the diagram of the phenakistoscope is from Helmholtz (1896), and the illustration of the viewer observing the reflected images is from Wade and Heller (1997). (Image © Nicholas Wade.)

optical displacements. One possible explanation is that Euclid was assuming that the eyes followed the moving objects. Support for this interpretation derives from the fact that the previous quotation is sandwiched between two others that are concerned with eye movements, and both define the perceived motion in terms of the optical geometry. For example:

> *If, when several objects move at unequal speed, the eye also moves in the same direction, some objects, moving with the same speed as the eye, will seem to stand still, others, moving more slowly, will seem to move in the opposite direction, and others, moving more quickly, will seem to move ahead.* (Burton, 1945, p. 371; italics in original)

Therefore, if the eye was following the two moving objects, the stationary one would move in the opposite direction with respect to the eye, and this would, according to Euclid's theory, account for the perceived motion.

There is no such doubt about the description given two centuries later by Lucretius (c. 98–55 BC; 1975) of a naturally occurring instance of induced motion: "when the winds carry scattered clouds across the sky in the night time, then the shining stars seem to glide against the clouds and to pass above them in a very different direction from their true one" (p. 311). A distinction is implied between the real and apparent motion of the stars, presumably because the stars appeared to remain stationary when the clouds were not visible. Similarly, Ptolemy noted the phenomenon with regard to water flowing past a stationary boat:

> When a boat is stationary on a calm river flowing without waves, a passenger who does not look at the distant land (but only at the water) thinks that the boat

> is moving quickly and that the water is still. . . . Since the water appears stationary, the motion must be seen as coming from the boat. However, if the water, land, and boat are observed at the same time, and if we are aware that the land is stationary, we will see the water in motion. (Lejeune, 1956, pp. 78–79; see Smith, 1996, for an alternative translation)

Ptolemy's account displays an appreciation of the relativity of motion perception with respect to the frames of reference that are in operation.

Perhaps the most common and compelling instance of induced motion is when moving clouds partially obscure the moon rather than stars, as Ibn al-Haytham noted: "Sight therefore perceives the moon as moving behind the swiftly moving cloud because it correlates a part of the moon with successive parts of the cloud" (Sabra, 1989, p. 261). This observation has been repeated and reported many times in the past millennium, as it provides one of the most readily visible instances of induced motion in the natural environment. For example, William Porterfield (1696–1771; 1759) described the natural occurring instance and gave a generalization from it (Fig. I.1-11):

> If two or more Objects move with the same Velocity, and a third remain at rest, the Moveables will appear fixed, and the Quiescent in Motion the contrary Way. Thus, Clouds moving very swiftly, their Parts seem to preserve their Situation, and the Moon to move the contrary Way. (p. 424)

Such examples of induced motion were discussed generally in 18th-century books on vision. For example, a clear link between such perceived motion and unconscious eye movements was made by Joseph Harris (1702–1764; 1775):

> If two or more objects, having the same apparent velocity, move all the same way; an object at rest, by which they pass, may appear to move the contrary way, whilst the objects in motion may seem at rest. For no motion among themselves is perceivable, since their images upon the retina keep the same distances from each other; and if the spectator insensibly moves his eye, so as to keep these images in the same place, the image of the object at rest will pass successively over them, after the same manner as if that object had been in motion the contrary way. (p. 106)

Thus, after 2,000 years, Harris adopted much the same interpretation of induced motion as Euclid. The additional feature that is made explicit in Harris's account is that the pursuit movements were insensible or unconscious. Harris (1775)

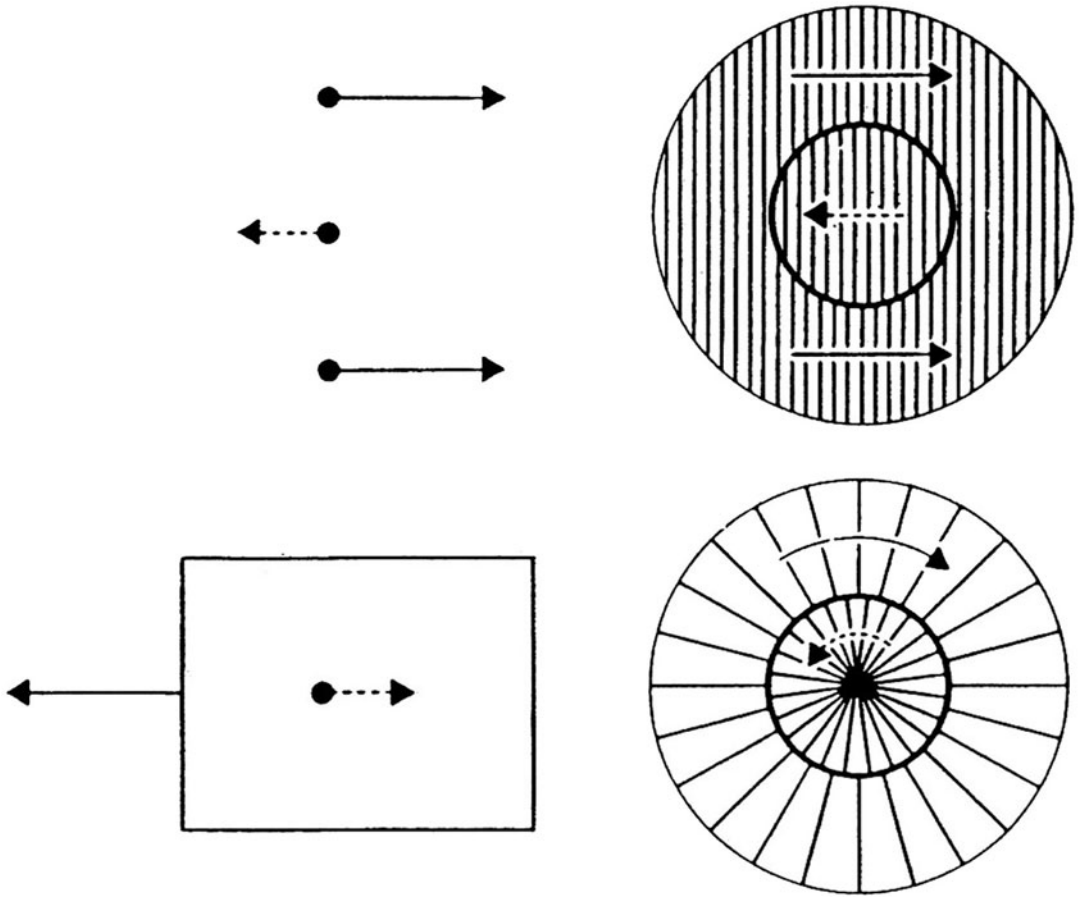


Figure I.1-11. Stimulus arrangements that have been employed to study induced motion, redrawn after original descriptions. Solid lines denote physical motion, dashed lines apparent motion. When two points move and a third remains stationary, the latter appears to move in the opposite direction (corresponding to Porterfield, 1759). Movement of a surround gratings results in the stationary central gratings apparently moving in the opposite direction (after Szily, 1905). Movement of a frame around a stationary dot was examined by Ehrenstein (1925) and by Duncker (1929), who also investigated rotation of an annulus around a stationary radiating center (after Wade and Swanston, 2013). (Image © Nicholas Wade.)

had applied the concept of unconscious inference in the context of binocular single vision:

> When both eyes look at the same object, the mind sees that object single. And it is very certain, that in at least abundant of cases, the visive faculty is assisted by a kind of reasoning, drawn from former experience, tho' perhaps unattended at the time we use it. (p. 111)

The concept of unconscious inference was to have wide currency in the 19th century: it was adopted and amplified by Wheatstone (1838) in the context of stereoscopic vision, and it provided a cornerstone for Hermann Helmholtz's (1821–1894; 1867/2000) theory of vision (Finger & Wade, 2002).

Experimental studies of induced motion commenced rather late, considering this rich history. Induced motion was referred to as motion contrast by German researchers in the early 20th century, and it was often linked to the MAE. For example, Adolf von Szily (1847–1920; 1905) fixated a point in the center of a stationary disc while a surrounding pattern of striped lines moved laterally—the disc appeared to move in the opposite direction, which was motion contrast (Fig. I.1-11). However, when the disc was viewed subsequently, relative to a stationary background, it appeared to move in the same direction as the background had formerly. This observation was confirmed by Basler (1910), who added many subtle variations to stimulus and observation conditions.

Motion can be induced with a simple arrangement consisting of a pair of moving and a single stationary stimulus (usually dots), as was described by Porterfield (1759). This was one of the situations examined by Duncker (1929); under these conditions the physically stationary dot is seen as moving in the opposite direction. A more compelling stimulus for inducing motion in a stationary dot is to displace a surrounding frame. Ehrenstein (1925) presented a diamond-shaped frame in two successive positions with respect to an enclosed stationary dot; the dot appeared to move in the opposite direction to the displacement (Fig. I.1-11). Duncker also used a moving rectangular frame surrounding a stationary dot. Under these circumstances, the rectangle can be moved below its detection threshold and still induce motion in the dot. The possibility that the inducing motion resulted in unconscious eye movements was still entertained, and both psychophysical and physiological tests of this hypothesis were examined. In addition, Duncker described the rotation induced in a stationary radial pattern by rotation of a patterned annulus or surround (Fig. I.1-11). The eye does undergo torsion as a consequence of surround rotation, but it is of limited extent and so is unlikely to be the basis of induced rotation. Evidence from measures of eye torsion confirms this conclusion—stimuli that modulate the degree of torsion that can be measured do not influence the magnitude of induced rotation reported (Wade, Swanston, Howard, Ono, and Shen, 1991). Another psychophysical technique involves inducing motion in opposite directions simultaneously (Gogel, 1977); it is difficult to argue that the eyes can move in opposite directions simultaneously. Finally, direct measures of eye movements and linear induced motion fail to demonstrate their dependence (Levi & Schor, 1984).

A more subtle version of the eye movement hypothesis is that it is either uncompensated eye movement signals or tendencies for the eyes to move that are correlated with the induced visual movement. These have the same problems in dealing with apparent rotations, motions in depth, and opposite linear motions as do simpler eye movement interpretations. Accordingly, the perception of motion in a stationary stimulus induced by a larger moving one is unlikely to be explained by any correlated movements of the eyes.

Motion Aftereffect

Induced motion is an example of simultaneous motion contrast, whereas the MAE reflects successive motion contrast. Having observed a moving stimulus for some time, a stationary one appears to move in the opposite direction. It can readily be seen after viewing descending water, and it was called the waterfall illusion in the 19th century; it was a motion illusion that excited considerable experimental and theoretical attention. Aristotle's description was followed by a more detailed account by Lucretius, also in the context of flowing water (Wade, 1994; Wade & Verstraten, 1998). Unlike most other phenomena described in antiquity, this motion illusion disappeared from view: there do not seem to be any other descriptions of it until early in the 19th century when it was rediscovered frequently. The two most notable accounts were made by Purkinje (1820, 1825) and Robert Addams (c. 1800–1875; 1834). Purkinje (1820) first briefly described the MAE in an article concerned with vertigo:

> Another form of eye dizziness can be demonstrated if one observes a passing sequence of spatially distinct objects for a long time, e.g. a long parade of cavalry, overlapping waves, the spokes of a wheel that is not rotating too fast. When the actual movement of the objects stops there is a similar apparent motion in the opposite direction. (pp. 96–97)

Purkinje (1825) amplified one aspect of the MAE in his second book on subjective visual phenomena. This report is more frequently cited: "One time I observed a cavalry parade for more than an hour, and then when the parade had passed, the houses directly opposite appeared to me to move in the reversed direction to the parade" (p. 60). Purkinje specified the direction of apparent motion, and he interpreted it in terms of involuntary eye movements over the houses. A few years later Addams (1834) observed what later became called the waterfall illusion at the Falls of Foyers in northern Scotland (Fig. I.1-12). Addams described the waterfall illusion thus:

> Having steadfastly looked for a few seconds at a particular part of the cascade, admiring the confluence and descussation of the currents forming the liquid drapery of waters, and then suddenly directed my eyes to the left, to observe the vertical face of the sombre age-worn rocks immediately contiguous to the water-fall, I saw the rocky face as if in motion upwards, and with an apparent velocity equal to that of the descending water. (p. 373)

The full text of Addams's brief but insightful article can be found in Dember (1964) and in Swanston and Wade (1994).

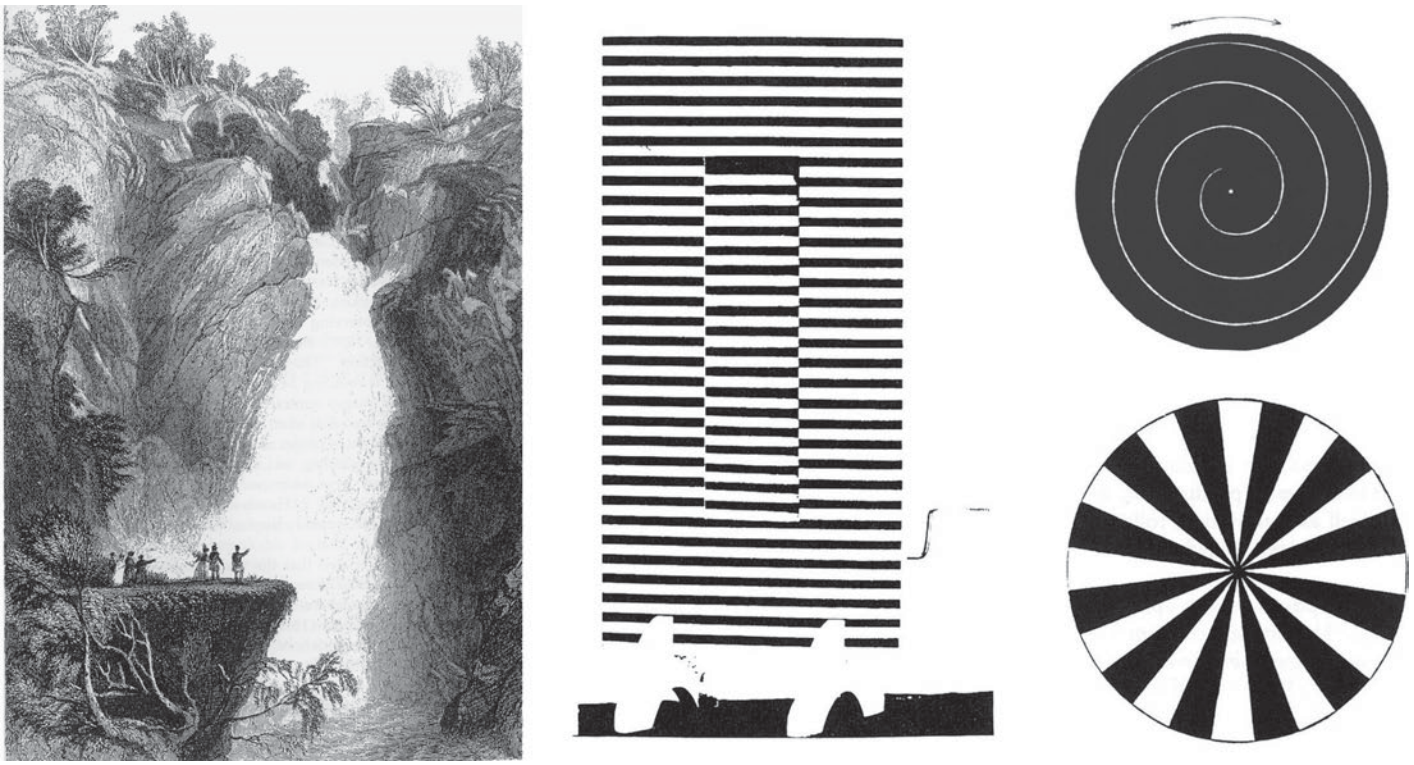


Figure I.1-12. Stimuli for investigating motion aftereffects. Left: A 19th-century engraving of the Falls of Foyers (Beattie, 1838); Addams would have observed the waterfall from the observation platform shown. Center: An artificial waterfall of the type used by Bowditch and Hall (1881). Upper right: Plateau's (1849) spiral. Lower right: A sectored disc as illustrated by Aitken (1878). (Image © Nicholas Wade.)

Despite the fact that Purkinje's book was well known in Germany, and that Addams' article was translated into German (Addams, 1835), both were often overlooked by German sensory physiologists, and the MAE was rediscovered independently several more times. Some of the rediscoveries were in the context of flowing water (Aitken, 1878; Müller, 1838; Oppel, 1856). Traveling on the railways also provided a platform for observing MAEs (Brewster, 1845; Helmholtz, 1867/1925; Thompson, 1877), and Thompson extended the MAE to depth as well as direction. MAEs in the third dimension were examined more systematically by Sigmund Exner (1846–1926; 1888) and Szily (1905). Thompson (1880) did make reference to Addams' report and was probably the first to refer to the phenomenon as the waterfall illusion (p. 294).

Addams (1834) appreciated that the waterfall illusion could be investigated experimentally, and he suggested that the motion of falling water could be simulated in the laboratory by moving stripes, but this was not put into practice until Johann Joseph Oppel (1815–1894; 1856) reached the same conclusion independently. Oppel had experienced the MAE initially at the dramatic Rheinfall at Schaffhausen, and this stimulated his interest in it. A few years earlier, Plateau (1849, 1850) stumbled across the MAE in the course of conducting experiments using the phenakistoscope (or stroboscopic disc). He had been studying the effects of rotating patterns on perception and noticed the motion visible in stationary objects following prolonged exposure. He introduced the stimulus employed most widely throughout the second half of the 19th century—a black disc with a white Archimedes spiral on it—which became called the Plateau spiral (Fig. I.1-12). Plateau (1849) described the MAE in the following way:

> If the disc rotates in the direction indicated by the arrow, and one looks at it with the eyes fixed on the centre for a time sufficiently long, but not long enough to tire the eyes, then one immediately directs the eyes to another object, such as the face of a person, for example, one experiences a singular effect: the head of the person appears to shrink for some time. If the disc is turned in the opposite direction, the resulting effect is opposite: it is as if the head of the person appears to expand. (p. 257)

The description is all the more remarkable because Plateau was blind when he wrote it (Verriest, 1990). Many of the 19th-century studies of MAEs used the Plateau spiral, or some variant of it, like concentric counter-rotating spirals (Dvorak, 1870). Other stimuli were introduced, however, some of which have a remarkably contemporary ring to them. Oppel (1856) constructed an instrument for producing continuous linear motion of parallel stripes. Bowditch and Hall (1881) added a further modification to the moving stripes, surrounding them with stationary ones (Fig. I.1-12). It was with such a stimulus that Wohlgemuth (1911) carried out many of his experiments.

Linear motion in two directions simultaneously was examined by Exner (1887) and by Borschke and Hescheles (1902). Horizontal and vertical gratings were moved vertically and horizontally, respectively, behind a circular aperture. Exner reported that the motion seen during adaptation was in a diagonal direction, with an MAE in

the opposite diagonal. Borschke and Hescheles developed the stimulus further with sets of vertical and horizontal rods: they systematically changed the relative velocity of the two components, thereby modifying the direction of the resultant vector and also of the MAE.

The sectored disc (Fig. I.1-12) was described independently by Wundt (1874) and by Aitken (1878), although it was by no means as widely used as Plateau's spiral. Rotating sectored discs continue to be used in experiments, and they are particularly good for demonstrating the paradoxical aspect of the MAE: following inspection the stationary sectors appear to rotate but not to change position.

The earliest interpretations of the waterfall illusion were in terms of eye movements (Addams, 1834; Helmholtz, 1867/1925; Purkinje, 1820) despite the logical difficulty of accounting for the phenomenon in this way: any aftermovement of the eyes would affect all contours in the visual field, rather than being restricted to regions exposed to prior motion. Mach (1875) pointed to the problems associated with eye movement interpretations and the difficulty they would have in accounting for radial motion seen in the spiral MAE. Kleiner (1878) rotated a central sectored disc in the opposite direction to two adjacent ones and reported oppositely directed MAEs that were visible simultaneously. Modern theories of the MAE are given in Mather, Verstraten, and Anstis (1998).

Visual Vertigo

The early-19th-century studies of MAEs were stimulated by the investigations of visual vertigo, although the phenomenon had long been known. The dizziness and giddiness caused by alcohol, looking down from heights, as well as a variety of ailments were known in antiquity (Brandt, 2003; Wade, 2000). There were also descriptions of the visual effects of body rotation: Lucretius described it as a children's game, and Ptolemy related vertigo to prior exposure to visual motion.

Reports of visual vertigo, like those of induced motion, have a more or less continuous history since antiquity. This probably reflects the fact that the experience is a common accompaniment of disease and so has retained an interest for medical reasons. The early modern era of research on vertigo began with Willis (1672), who both defined it and interpreted it in mechanistic terms. Vertigo was considered to involve motion of the animal spirit in the head, which produced apparent motion during rotation, rather like smoke in a flask lagging behind that of the rotating vessel. Moreover, Willis described the visual motion that continues after body rotation ceases, and this was attributed to the continued motions of the animal spirit relative to the stationary head:

> Vertigo arises from the circular motion of the spirits, and, as it were, their rotations in the brain and its medullary part. It takes place just as smoke and vapour contained in a glass or phial are sent into similar motion if you spin the vessel round. This motion lasts longer in the smoke or vapour than in the vessel. Thus we find people whose spirits are very thin, and therefore flexible and weak, pass into vertigo as soon as the body or head is rotated and this sensation persists after the body has ceased its turning motion. (Dewhurst, 1980, pp. 113–114)

Research on postrotational visual motion commenced in earnest in the 18th century, largely as a consequence of Porterfield's (1759) association of eye movements with visual motion. Motion was the last of the phenomena of vision described in the second volume of Porterfield's *Treatise on the Manner and Phœnomena of Vision*, and his analysis of it was subtle: "Our eyes do never inform us of what Magnitude or Celerity of Motion is, considered absolutely and in itself, but only what it is relatively, when compared to some other Motion" (p. 417). He went on to list a range of motion phenomena, including the lucid definition of visually induced motion described previously. Vertigo was the final visual phenomenon discussed, and he argued that the eyes did not move when the body stopped rotating. The evidence was subjective: Porterfield was not consciously aware of any movements of his own eyes, and so he was convinced that they remained stationary following rotation (Wade, 2003, 2007c).

Eye movements interact with afterimages, which were known to move with voluntary movements of the eyes. It was this aspect that attracted both Erasmus Darwin (1731–1802) and his son, Robert Darwin (1766–1848), to Porterfield's account of postrotational visual motion. Robert Darwin (1786) presented a detailed account of afterimages (which he referred to as spectra), distinguishing between their positive and negative phases, relating their visibility to differential retinal adaptation, and describing how their apparent size was related to the distance of the surface onto which they were projected. In addition, he offered an alternative interpretation of postrotational apparent motion. He suggested that the sequence of light entering the eyes during body rotation was visible as a moving afterimage when the body came to rest. Thus, Robert Darwin suggested that visual vertigo was a consequence of moving afterimages, a view encouraged by his father, who, it has been suggested, might have had a substantial hand in writing his son's article (King-Hele, 1999, 2002; Wade, 2002).

Neither Erasmus nor Robert Darwin accepted that there was any link between eye movements and visual vertigo. Erasmus Darwin (1794) maintained that: "The cause of thus turning our eyes forwards, and then back again, after our body is at rest, depends, I imagine, on the same circumstance, which induces us to follow the indistinct spectra" (p. 571). That is, there is a tendency to pursue but never to locate a peripheral afterimage even when it is very weak. Erasmus Darwin did conduct an experiment that provided what he considered to be crucial evidence that eye and visual motions were not related. He denied the occurrence of ocular torsion or rolling of the eyes, and so he conducted an experiment:

> in which the rolling of the eyes does not take place at all after revolving, and yet the vertigo is more distressing than in the situations above mentioned. If any one looks steadily at a spot in the ceiling over his head, or indeed at his own finger held up high over his head, and in that situation turns round till he becomes giddy; and then stops, and looks horizontally; he now finds, that the apparent rotation of objects is from above downwards, or from below upwards; that is, the apparent circulation of objects is now vertical instead of horizontal, making part of a circle round the axis of the eye; and this without any rolling of his eyeballs. The reason of there being no

> rolling of the eyeballs perceived after this experiment, is, because the images of objects are formed in rotation round the axis of the eye, and not from one side to the other of the axis of it; so that, as the eyeball has not the power to turn in its socket round its own axis, it cannot follow the apparent motions of these evanescent spectra, either before or after the body is at rest. (pp. 570–571)

Thus, Erasmus Darwin considered that if ocular torsion was not possible, then rotary motion could not be associated with it! Accordingly, Darwin dismissed the correlation between eye movements and visual motion for all forms of postrotational vertigo.

The modern era was heralded by a much-neglected American scientist of Scottish descent, William Charles Wells (17571817; 1792, 1794a, 1794b). Although born in Charlestown, South Carolina, he was educated mainly in Scotland and practiced medicine in London (Wade, 2003). Wells (1794b) engaged in a public dispute with Erasmus Darwin over the interpretation of visual vertigo and summarized his conclusions:

> When we stop ourselves while giddy from turning, our eyes do not return to the state of rest along with our bodies, but continue to move for some time after. Of this, however, we are not conscious; and hence we imagine the relative motion between our eyes and objects at rest to be possessed by the latter. (p. 905)

The elegance of Wells's experiments lies not only in the description of the phenomenon but also in the use of afterimages, generated before rotation, to render visible the motions of the eyes after rotation ceased (Tatler & Wade, 2003; Wade, 2000, 2010b, 2010c). Wells can be considered to have initiated the modern study of eye movements by seeking a more precise method of measuring them, namely by using afterimages as stabilized retinal images. He described the phases of nystagmus and the suppression of vision during the fast phase; he related the direction of eye movements to that of prior rotation; and he demonstrated torsional nystagmus under the conditions described by Erasmus Darwin. Moreover, Wells explicitly related the movements of the eyes to those of the visual world following body rotation.

Mach (1875) provided support for Wells's conclusions, although Wells was not mentioned by him. Mach also generated an afterimage prior to rotating his body and noted its apparent motion when rotation ceased (see Young, Henn, & Scherberger, 2001, for a translation of Mach's work on the movement sense). Breuer (1874) obtained a similar result, and Crum Brown (1878) gave graphic depictions of eye movements during and following body rotation. As would be expected of a phenomenon that is related to eye movements, visual vertigo involves a rotation of the whole visual world, rather than a part of it. An illustrated account of the pioneers of eye movement research can be found in Wade (2010b).

Autokinetic Effect

Yet another motion phenomenon relates to the apparent movement of stars at night. In 1799, the great German naturalist Alexander von Humboldt (1769–1859; 1850) recorded that when he observed certain stars, they appeared to wander about haphazardly in the night sky:

> Here is the place, at least in passing, to mention another optical phenomenon which I observed only once in all my mountain climbing, which occurred before sunrise on 22 June 1799 on the slopes of the Pics of Tenerife. In the Malpays, at a height of 10,700 feet above sea level, I saw with my naked eye low-lying stars in a wonderful wandering motion. Illuminated points rose upwards, moved sideways and fell back to their original positions. The phenomenon lasted 7 to 8 minutes and came to a halt long before the appearance of the sun above the horizon. The same motion could be seen through a telescope; there remains no doubt that it was the stars themselves that moved. (p. 73)

The observation was made during a scientific voyage to South America, and it is clear that Humboldt genuinely believed that the stars did undergo the rapid motions that he perceived. He even called the phenomenon *Sternschwanken* or star wandering. Exner (1896) remarked that the same observation was made from the same spot independently some 50 years later by Prince Adalbert von Preussen. Other observers were able to see the phenomenon too, though they did not all describe the haphazard movements in a similar way. A debate arose about the precise basis for the star movements, and its subjective nature was determined by Schweizer (1857, 1858), an astronomer in Moscow; his work has been summarized by Exner (1896) and Holt-Hansen (1962). Schweizer (1857) made many observations on the phenomenon, both under naturalistic conditions and in the laboratory. His evidence that the visual motion was subjective rather than objective derived from a simple study in which he and a colleague observed Sirius together and compared their reports of the motion paths:

> By this means a new fact was established that *at one and the same moment* the appearance was different for each of us. Sometimes I saw the star rise while Herr Bredichin saw it fall; at times it appeared stationary to me while the other observer caught sight of it in rapidly advancing movement; at yet other times the star described a circle for me while for Herr Bredichin it rose or fell in a slanting direction. (p. 443)

As there was no consistency in the motions of the same star seen at the same time, it became clear that it was a perceptual rather than a physical phenomenon. The phenomenon is not confined to stars in the sky, and Schweizer simulated the conditions with artificial stars in a darkened room. If one looks at a stationary and very dim light in an otherwise totally dark room, it will appear to wander about haphazardly, like the stars that Humboldt observed. Schweizer noted that the movements in a horizontal direction were more pronounced than those in the vertical. He also observed the motion with a black spot on a white wall, and vice versa; after a short time the spot moved in various directions but always returned to the original point. This was referred to as *Punktschwanken* to distinguish it from *Sternschwanken. Scheinbewegung* or apparent motion was applied to the movement of a faint light in a dark room (Exner, 1896). Adams (1912) suggested that they should all be included under the rubric autokinetic sensations.

Most early reports noted that there is a latency of some seconds before the autokinetic motion is seen, and thereafter the amplitudes can be considerable. There was, however,

much debate about the pattern and velocity of the motion seen. Charpentier (1886) described the motion as radial and slightly curved with a low velocity; the amplitude could be over 30°. Carr (1910), on the other hand, reported slow and irregular motions in the central visual field, but the velocity was estimated as up to 20°/s when the eyes were fatigued.

Hermann Aubert (1826–1892) was initially unaware of Charpentier's experiments when he independently rediscovered autokinetic motion in 1887, and both failed to cite Schweizer's studies. Some years earlier, Aubert (1865) imposed some degree of order on a wide range of visual phenomena, particularly those involving motion. For example, he measured the thresholds for detecting uniform and relational motions. When a target moved in isolation, its motion could be detected when it reached a velocity of about 0.3°/s. When a patterned background was visible, the threshold was approximately 0.03°/s. Since isolated stimuli have higher thresholds for motion, they are only seen to move when they undergo quite large displacements. Thus, there is clearly an influence of the visible reference frame in motion perception, but that was not considered to be the sole factor. Aubert returned to the examination of motion phenomena in the 1880s and rediscovered the apparent motion of a static light in a dark room; he gave it the name autokinetic sensation (Aubert, 1887).

As with the phenomena mentioned previously, eye movement explanations were the first candidates for the autokinetic sensations, but experimental investigations failed to find a simple link. Carr (1910) adopted an afterimage technique to examine this by comparing the movements of the afterimage with those of a spot of light. The afterimage remained relatively stable, even during quite large apparent movements of the spot. Using direct recordings of eye movements, Guilford and Dallenbach (1928) found that their direction and amplitude did not correlate with those of autokinetic motion. However, there were some aspects of the phenomenon that could be related to the direction of eye movements. For example, autokinetic motion is in the same direction as prior eccentric fixation (Carr, 1910), and it can be related to the physiological position of rest of an eye (Levy, 1972). Nystagmic eye movements have been induced by rotary and caloric stimulation, and these produce similar apparent motions of an isolated light (Skolnick, 1940). As with visual vertigo, the apparent motion was associated with involuntary rather than voluntary eye movements. In the autokinetic effect there is no stimulus motion, so the perceived movement must be caused by some internal process, probably related to actual or suppressed movements of the eyes (Levy, 1972; Post & Leibowitz, 1985). An assessment of these interpretations can be found in Howard (1982) and Mack (1986).

Eye movements are such an obvious aspect of vision that they have always been of interest. They were remarked upon in antiquity, but determining the characteristics of such eye movements is a more recent preoccupation. Most specifically, their measurement has been rendered independent of the observer. That is, the methods that were used for measuring eye movements in the past (like placing a finger over the eyelid, or observing the eyes of another person) have been improved by more modern techniques. Initially the use of afterimages enabled displacements of the stabilized from unstabilized images to be noted. This method was adopted widely in the 19th century (Tatler & Wade, 2003; Wade & Tatler, 2005, 2011). Technical improvements, like tracking Purkinje images or the vessels in the iris, have added precision to the determination of eye movements and eye position (Heller, 1988).

The distinguishing feature between visual vertigo and the phenomena of induced motion, MAEs, and stroboscopic motion is that the first involves motion of the whole visual field, whereas for the other three motion is experienced in part of the visual display. That is, the whole scene rotates during visual vertigo rather than any enclosed parts of it. The importance of relative motion has been demonstrated in a MAE that is produced by induced motion (Swanston & Wade, 1992). The autokinetic effect does not involve relativities of motion, as the light that appears to move is observed in isolation. The search for eye movement correlates of motion illusions in which some objects appear to move relative to others has not proved successful. It reflects a desire to link an objective index (eye movements) to a subjective experience (apparent motion). In the absence of such a link, visual neuroscience is faced with relating subjective experience with measures of neural functioning (Wade & Heller, 2003). It should resist the temptation of denigrating the subjective at the expense of sophisticated measures of neural activity. If one of the aims of visual science is to relate subjective experience to objective indices, then sight should not be lost of the subjectivity of the experience. It is to subjective visual phenomena that we now turn.

SUBJECTIVE PHENOMENA AND OBJECTIVE CORRELATES

The subjective dimension of vision tends to be treated as an oddity rather than a fundamental feature of perception. In many ways, the development of visual science has reflected a rejection of the subjective experience in human behavior. The behaviorist revolution was successful because it concentrated on those aspects that were considered as objective—stimulus and response. Perception has always proved difficult to constrain in this manner. For example, in the early 19th century, certain phenomena were referred to as subjective, as there was little control of the stimuli that gave rise to them. This resulted from Purkinje's (1823, 1825) description of a range of such effects (Wade & Brožek, 2001). However, even at this stage objections to the use of the term were provided. In his doctoral dissertation, Purkinje (1823) described subjective phenomena in these words:

> There are sensations which do not correspond to anything outside the body. In so far as they imitate the qualities and forms of external things, they thereby often give rise to illusions, phantoms, or appearances with no corresponding reality. These can be referred to as subjective sensory phenomena. (pp. 3–4)

His illustrations of the phenomena are shown in Figure I.1-13.

Subjective visual phenomena had a long history before Purkinje gave them that title, but he added greatly to the detail of their classification, description, and interpretation. The description of one's own sensations might seem the simplest of things to report upon, but this is not the case. Helmholtz (1867) appreciated this only too well, and he commended Purkinje's special talent in this regard.

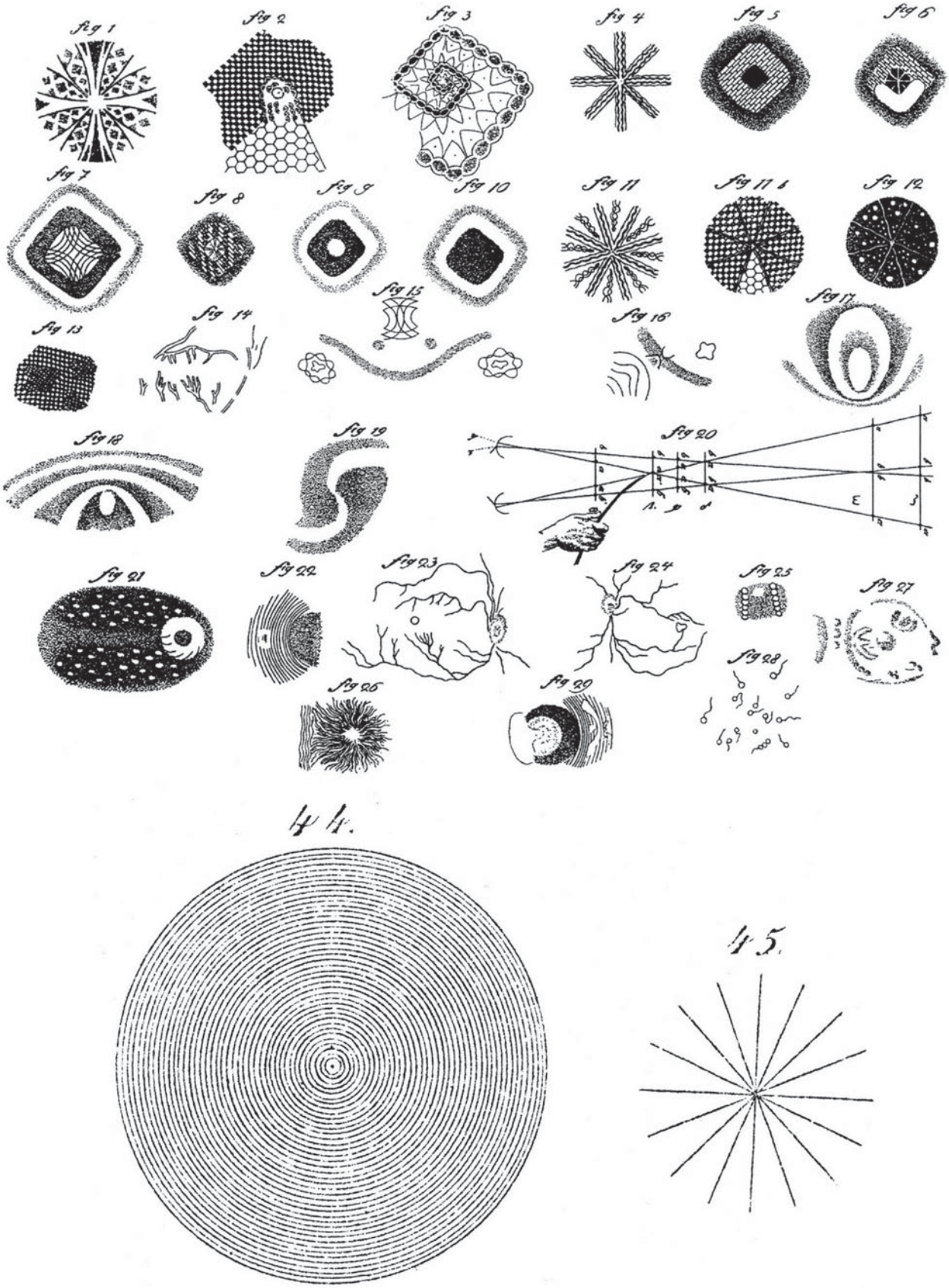


Figure I.1-13. Upper: The plate from Purkinje's (1823) book on subjective visual phenomena. Figures 1-4 were light and shadow figures; 5–14 were patterns visible when pressure was applied to the eye; 15–16 were experiences seen when the eye was stimulated electrically; 17–19 were cloudy streak patterns; 20–22 involved locating the blind spot and representing the haloes seen around it; 23–24 were the shadows of retinal blood vessels cast by light at the corner of the eye (this pattern is known as the Purkinje tree); 25–29 represent the visibilities of entoptic phenomena that Purkinje called blood circulation in the eye, curvilinear radiations, pulsating figures, flying gnats, and luminous rings, respectively. Lower: Two figures from Purkinje (1825) showing patterns of concentric circles and radiating lines; these appear to become distorted and move with prolonged viewing and they also generate dramatic aftereffects when a blank surface is subsequently viewed. (Image © Nicholas Wade.)

Purkinje commenced his research on visual phenomena because he had little access to equipment that would have allowed him to conduct experimental enquiries into other aspects of physiology.

Purkinje (1825) is perhaps best known for the brightness changes of colors at dawn or dusk: blue objects that appear brighter than red ones before sunrise reverse thereafter. This phenomenon is now called the Purkinje shift, and later in the century it was related to the different spectral sensitivities of rod and cone receptors in the retina. He described it thus:

> The degree of objective illumination has a great influence on the intensity of colour quality. In order to prove this most vividly, take some colours before daybreak, when it begins slowly to get lighter. Initially one sees only black and grey. Then the brightest colours, red and green, appear darkest. Yellow cannot be distinguished from a rosy red. Blue looks to me the most noticeable. Nuances of red, which otherwise burn brightest in daylight, namely carmine, cinnabar and orange show themselves as darkest, in contrast to their average brightness. Green appears more bluish, and its yellow tint develops with the increasing daylight. (pp. 225–226)

The Purkinje shift was described in the second book on subjective vision, which was published in 1825. Eighteen topics were covered in the book; some extended descriptions of phenomena addressed in the first volume, whereas others were novel. The principal new contributions were indirect vision, real and apparent movements, investigations of the interaction of colors, focal image inside the eye, visual flicker after the use of digitalis, some comments on distant and near vision, intentional squinting, and the effect of belladonna on vision. The description of the brightness changes of red and blue objects at dawn or dusk (the Purkinje shift) was given in the section on the interaction of colors. He also described a perimeter and color zones of the eye (under indirect vision), the MAE and visual vertigo (under real and apparent movement), the distortions visible in regular geometrical patterns, and the failure of accommodation following application of the mydriatic, belladonna. The color zones were charted with the aid of a perimeter. He found that colors could not be distinguished in the peripheral retina, that yellow and blue were recognizable at slightly greater peripheral angles than red and green, and that all colors were visible more peripherally in the temporal than the nasal fields. Prolonged observation of a pattern of concentric circles (shown in Fig. I.1-13) had the effect that

> there appear in all directions bands of clearly distinguishable parallel lines, over which the multitude of lines slide and entwine as cloudy streaks and points; they all radiate from the centre to the periphery, and their number, width and direction differ with different individuals, but remain constant for any one. (Purkinje, 1825, p. 261)

Johann Wolfgang Goethe (1749–1832; 1824) wrote an extensive review of Purkinje's first book. He reprinted particular short passages from it and then added his own comments in parentheses. These were either of a general nature or related to Goethe's own observations. He was clearly impressed by the subtlety of Purkinje's vision because he wrote "we are grateful to the author for undertaking this task and for raising it to a new level" (p. 103). Goethe was not alone in his admiration of Purkinje's powers of observation. In 1826, Johannes Müller (1801–1858; 1826a, 1826b) extended the work in two books titled *On Fantastic Visual Appearances* and *On the Comparative Physiology of Vision in Humans and Animals*. Later, in his influential text, *Elements of Physiology*, he remarked of these phenomena, "knowledge of which we are principally indebted to Purkinje" (1843/2003, p. 1210) before summarizing his own observations. In the 1830s, Purkinje and Müller were to found the two most influential laboratories of physiology—in Breslau and Berlin, respectively.

One of the few reviews in English of Purkinje's first book was under the initials C. W. (Charles Wheatstone); it appeared in 1830 and it was confined to Purkinje's (1823) first volume. Wheatstone commenced his review of Purkinje's book by noting that this little volume has excited considerable interest in Germany, but he took issue with the use of the term "subjective":

> To distinguish these phenomena from those which arise on the presence of their appropriate external objects, the author employs the term subjective, which, as denoting this class of phenomena better than any other we are acquainted with, and, to avoid circumlocution, we have purposely retained; it will, however, on consideration, be perceived, that the term is not strictly proper, as, correctly speaking, all phenomena, *as such*, are subjective, *i.e.* in the mind; and were we, without qualification, to admit the classification of phenomena into objective and subjective, we should be unable to determine, with any degree of accuracy, where the objective ends or the subjective begins. (p. 102)

Wheatstone's strictures apply as much today as they did then. He not only gave a summary of selected sections from Purkinje's book, but he also added novel methods for observing some phenomena—particularly the visibility of retinal blood vessels. Wheatstone had an important theoretical influence on Helmholtz, but this was mainly through his work on binocular vision. Helmholtz represented the shift from the subjective to the objective in German research on the senses. That is, he turned his back on the subjectivism championed by Goethe and the movement of *Naturphilosophie*, placing perception in the Newtonian mold. The struggle was a difficult one because Goethe was in the German rationalist philosophical tradition whereas Helmholtz sought support from British empiricist philosophy.

Purkinje opened a new domain of inquiry in the long history of vision research—that of subjective visual phenomena. His remarkable observations were lauded by both phenomenologists and empiricists. With regard to Purkinje's first book, Goethe noted:

> One has to be born strong to be able to examine his own interior without fear. The ability to take a healthy glance back into one's own interior without being buried by it, coupled with talent for profound investigation into the unexplored problems with a clear insight but not with illusion and fantasy, are precious gifts; the results of this research indeed are rare good luck for both the world and science. (John, 1959, p. 18)



Figure I.1-14. Early images of subjective contours. Upper left: From Helmholtz (1867). Upper right: From Ehrenstein (1941). Middle: Jastrow's (1899) incomplete letters. Lower: Separated semicircles from Schumann (1900). (Image © Nicholas Wade.)

After Purkinje's death certain subjective visual effects that he had not studied assumed prominence. These were mainly concerned with the appearance of contours where no luminance differences existed (Kanizsa, 1979). Surprisingly, a clear example of subjective contours was supplied by Helmholtz (1867) without an appreciation of the phenomenon. In fact, the figure was made to show variations in apparent length with orientation, and it was printed by the side of his vertical and horizontal squares (see Fig. I.1-16). Ehrenstein (1941) completed the circle started by Helmholtz but with the implicit recognition of subjective or illusory contours (see Lingelbach, 2012a). Ehrenstein also noted that "The brightness of the central area increases up to a certain point with increasing thickness of the rays, but also with increasing number of these rays" (p. 36). The impetus for Ehrenstein's investigation was not Helmholtz's figure, but the illusory dots seen in the Hermann-Hering grid (see Fig. I.1-18). Joseph Jastrow (1863–1944; 1899) was able to demonstrate the effect with letters (Fig. I.1-14). His concern was in proofreading and overlooking the spatial details of letters so that incomplete ones were seen as complete. A few years later, Friedrich Schumann (1863–1940; 1900) produced what he called subjective lines simply by separating the two halves of a circle (Fig. I.1-14). The articles by Schumann and Ehrenstein have been translated into English by Anne Hogg and published in Petry and Meyer (1987). Research on subjective contours exploded around 1970 (Purghé & Coren, 1992) and continues apace.

NINETEENTH-CENTURY ILLUSIONS

The 19th century witnessed an explosion of experimental ingenuity in all areas of science. The senses were at the center of many of the dramatic departures, and the experimental advances in turn influenced theories of perception generally

and illusions in particular. As an example, the stereoscope was invented in the 1830s; it transformed not only the vision of pictures but also the picture of vision. The hallmark of the century was the invention of instruments that enabled the experimental investigation of phenomena (Wade, 2004).

When Purkinje (1823) remarked that visual illusions revealed visual truths, he was not referring to geometrical optical illusions because they were not given this appellation until the second half of the 19th century (Phillips & Wade, 2015); this was the era of illusions. Despite the fact that illusions have fascinated students of the senses for over 2,000 years, many of the geometrical optical illusions graced by the names of their discoverers derive from this period (Robinson, 1972; Vicario, 2008, 2011; Wade, 2014). Helmholtz, writing in 1867, did describe a few illusions but not many. The avalanche descended in the later decades. Geometrical optical illusions are quintessentially phenomena of the late 19th century, when the likes of Mach, Müller-Lyer, Poggendorff, and Zöllner described their eponymous phenomena. These illusions have an important place in the history of psychology, because they were among the factors that led Wilhelm Maximilian Wundt (1832–1920) to establish his Psychological Institute at Leipzig in 1879—he could not envisage how illusions could be accounted for in physiological terms, and so they, along with consciousness, required a separate discipline. Wundt took as his yardstick the proximal stimulus (the retinal image)—and he could not accept that, say, two linear extents that produced equivalent retinal extents could yield perceptual inequality due to physiological processes. Thus, geometrical optical illusions are important in the context of establishing psychology as an independent discipline: there was considered to be no physiological correlate of perception. Psychologists like Wundt sought to determine correlates, and the one most favored at that time was in terms of eye movements, although other higher-level alternatives were also entertained.

It is, however, instructive to examine why this burst of illusory activity occurred in the late 19th century. In other words, why did outline drawings assume such a central role in the study of perception? It could have been due to the combination of two powerful strands of thinking about vision. The first stems from the 17th century, when Kepler (1604), Scheiner (1619), and later Descartes (1637) elucidated the dioptrical properties of the eye. They thereby set in motion the idea that the problem of perception has as its starting point the static retinal image. The problem was seen as restoring the missing dimension of distance from the ambiguous retinal projection. However, this is both a physical and a physiological fiction. The static retinal image is a convenient physical fiction because it allows us to draw ray diagrams that describe the dioptrical properties of the eye. It is a physiological fiction because the eye is never still and the receptors collect energy at differential rates. Nonetheless, these fictions continue to drive our models of vision and our analyses of illusions.

The second strand relates to the experimental approaches to the study of perception introduced in the mid-19th century. The physicists Wheatstone (1838) and Helmholtz (1867) argued that experimental rigor of the physical sciences should be brought to bear on the study of perception. Thus, stimulus variables should be isolated and manipulated in quite unnatural ways in order to determine how perception is modified. It is difficult to manipulate solid objects, but it is exceedingly easy to create novel pictures. Moreover, Wheatstone himself had shown that the perception of three-dimensional space can be synthesized from the use of two appropriate flat drawings. So pictures became the accepted stimuli for the study of vision. Once accepted, the psychologists then rediscovered tricks that had been a part of the artist's armoury for centuries, and they devised some novel ones too (Pinna & Reeves, 2009; Wade, 2007a, 2012b). Illusions were studied in the late 19th century because they were not amenable to the extant physiology—hence their place in psychology. They also fostered the use of two-dimensional stimuli in perceptual experiments.

GEOMETRICAL OPTICAL ILLUSIONS

The illusions that attracted the interests of many 19th-century visual scientists were those labeled "geometrical optical" by Oppel (1855). Oppel presented very simple drawings (Fig. I.1-15) in which single or parallel lines appeared to be bent by the superimposition of angled lines.

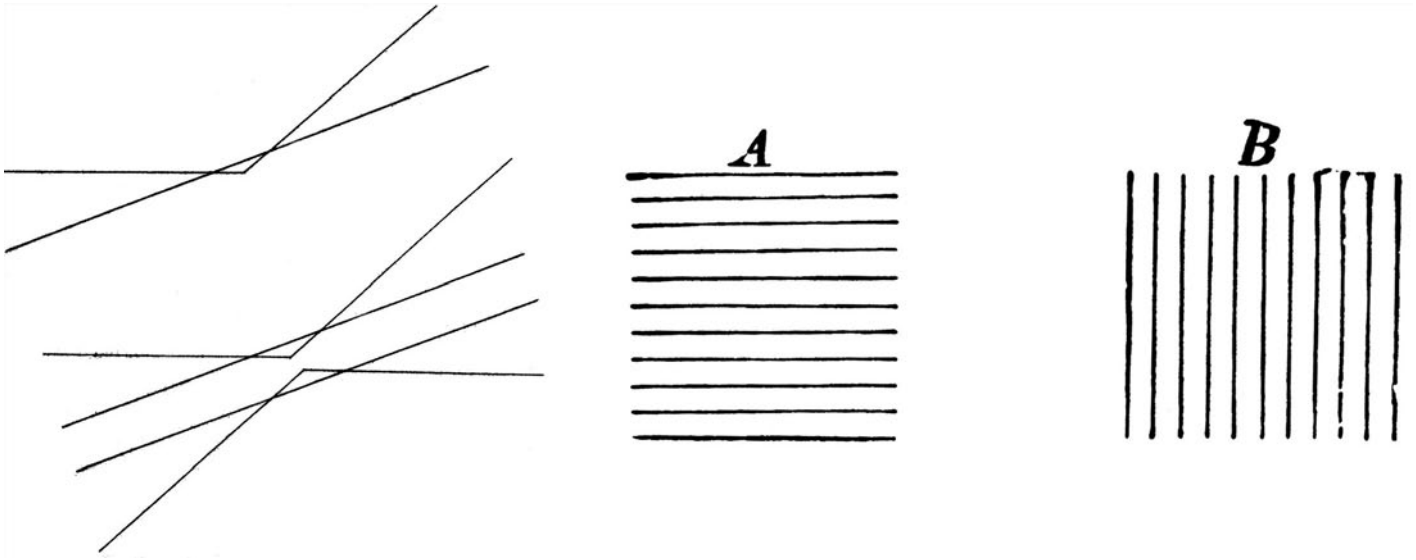


Figure I.1-15. Left: Figures 2 and 3 from Oppel's (1855) article. Right: Helmholtz squares illusion from Helmholtz (1867). (Image © Nicholas Wade.)

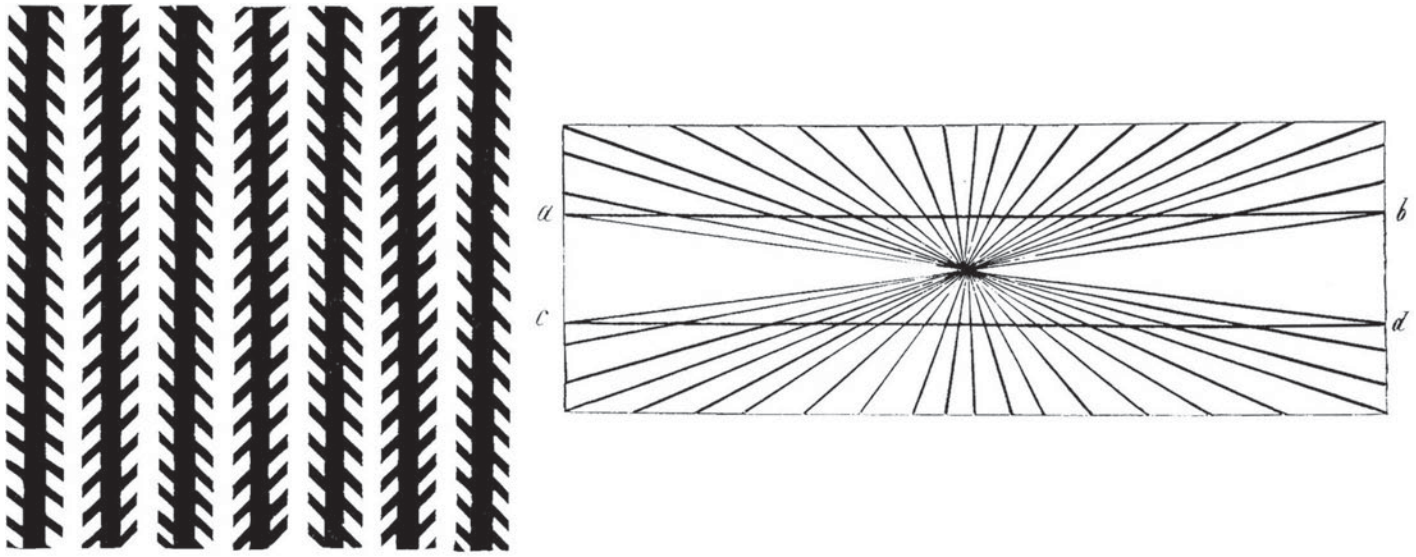


Figure I.1-16. Illusions of orientation. Left: Zöllner's (1860) illusion in which the vertical lines appear inclined in different directions due to the cross-hatched lines. Right: Hering's (1861) figure inducing curvature in the two horizontal lines because of the varying angles of intersection with the radiating lines. (Image © Nicholas Wade.)

Thus orientation illusions were the first to be enveloped by this title. Geometrical optical illusions consist of relatively small but reliable distortions of visual space, mostly in the domains of size or orientation. A configuration displaying distortions of both dimensions (the horizontal–vertical illusion) was described by the German physician, Adolf Fick (1829–1901; 1851) in the course of studying astigmatism. The illusion was considered commonplace by the time Oppel was writing, as he mentioned it without reference to Fick. Many novel forms were devised and depicted in the final decades of the century, and they often bear the names of those who first drew or drew attention to them (Robinson, 1972; Vicario, 2008, 2011).

The Helmholtz squares illusion (Fig. I.1-15) is based on the distortions of space in the cardinal orientations. It is related to what is now called the horizontal–vertical illusion, which is usually represented as single lines in an inverted T shape. It incorporates both orientation and extent, whereas most others involve one of these dimensions alone. Two orientation illusions are shown in Figure I.1-16. That on the left consists of the original figure used by Zöllner (1860) to show how cross-hatchings in opposite directions result in lines appearing to be inclined. Zöllner sent his figure to the *Annalen der Physik und Chemie*, which was edited by Johann Christian Poggendorff (1796–1877). When Poggendorff saw the figure in the manuscript copy, he noted that the edges of the cross-hatched lines appeared out of alignment. That is, the upper or lower edges of the oblique lines on one side of the thick verticals did not seem to be in alignment with those on the other side. It was pointed out to Zöllner, who acknowledged it in the article: "Hr. Prof. Poggendorff had the kindness to draw my attention to it. It is this, the Nonius-like displacement on both sides of the long stripes" (Zöllner, 1860, pp. 501–502). Ewald Hering (1834–1918; 1861) introduced a simple figure in which parallel lines appeared curved as a consequence of the radiating lines they crossed (Fig. I.1-16). Wundt (1898) presented a variant of Hering's illusion with the radiations centered at the extremities and meeting in the center; this resulted in parallel lines appearing curved in the opposite direction.

Oppel (1855) mentioned illusions of extent as well as orientation, and in a later article (Oppel, 1861) he went on to describe how divided or filled spaces appear larger than equivalent undivided separations. August Kundt (1839–1894; 1863) remarked on a similar effect, which has become known as the Oppel-Kundt illusion (see Wackermann, 2011). Although Oppel introduced some illusions of length, they were studied in detail somewhat later when they were subjected to more intense experiment and interpretation than those of orientation. A remarkable survey of this illusory scene was carried out a few years later by Joseph-Rémi-Léopold Delboeuf (1831–1896; 1865a, 1865b). The first article was concerned mainly with orientation illusions, but the second one presented many examples of irradiation (see Nicolas, 1995). However, the Delboeuf illusion is usually associated with two equal outline circles either surrounded by or enclosing another circle, which he illustrated in a later article (Delboeuf, 1892). The illusions of Franz Carl Müller-Lyer (1857–1916; 1889) and Mario Ponzo (1882–1960; 1912) are shown in Figure I.1-17. Many more geometrical optical illusions were described in the late 19th and early 20th centuries.

Contrasting elements can result in variations in apparent size, and one of these was described by several researchers, including Edward Bradford Titchener (1867–1927; 1901). It consists of two equal circles surrounded by larger and smaller ones. The illusion is also associated with Hermann Ebbinghaus (1850–1909), although it was not illustrated by him but by Armand Thiéry (1868–1955; 1896) who wrote: "A circle that is surrounded by several other circles appears larger when they are smaller than when they are larger (Fig. 51A u. B, Ebbinghaus)" (p. 84). Thiéry (1895a, 1895b) interpreted geometrical optical illusions in terms of perspective and provided examples of distortions based on perspective, as had von Bezold (1884) before him.

Brightness contrast phenomena provided another potent source of visual illusion. Although the influence

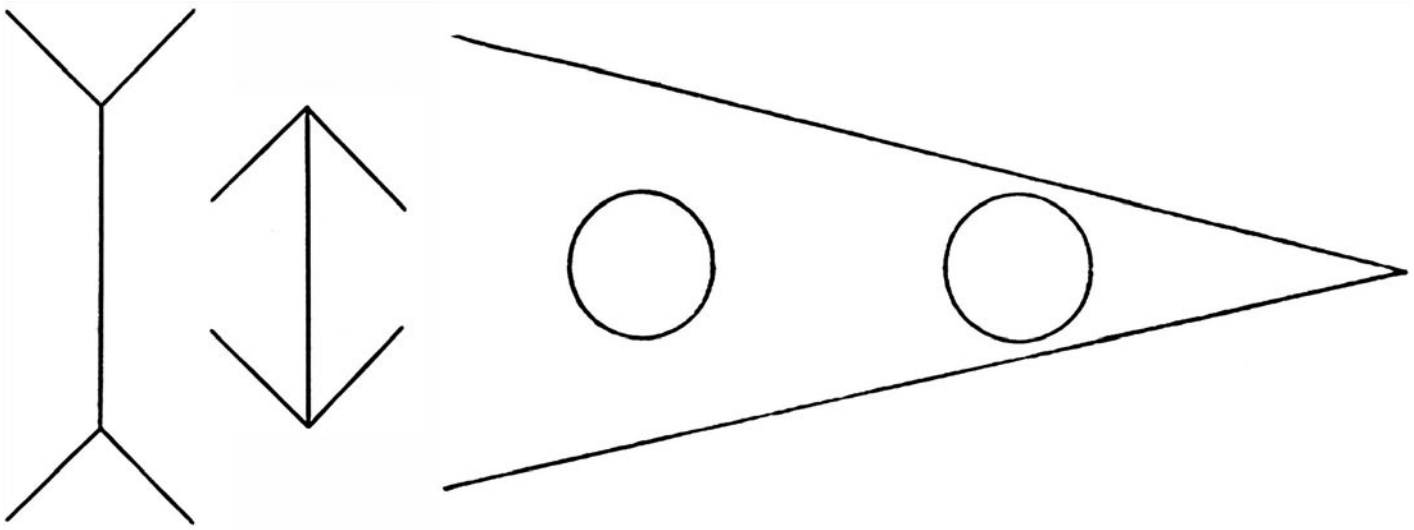


Figure I.1-17. Illusions of extent. Left: Müller-Lyer illusion as represented by him in 1889. Right: Ponzo illusion from Ponzo (1912). The vertical lines are the same length and the circles are the same size but they do not appear so. (Image © Nicholas Wade.)

that a surrounding color can have on an enclosed one had long been known, subtle effects of brightness contrast were described in the 19th century. For example, Ludimar Hermann (1838–1914; 1870) gave an account of the effects that can be seen in the grid on the left of Figure I.1-18—dark dots appear at the intersection of the lighter lines. He related the illusory dots to the operation of simultaneous contrast. The converse, where light grey dots are seen at the intersections of the darker lines, was described later by Hering (1907). Hermann noticed the effect when he viewed an illustration of the Chladni figures in a translation of Tyndall's (1869) book on sound (see Wade, 2005b).

The Hermann and Hering grids have assumed a position of considerable importance in contemporary visual science, as they have provided a means by which receptive fields recorded physiologically can be related to "perceptive fields" determined psychophysically (see Lingelbach & Ehrenstein, 2002; Schiller & Carvey, 2005; Spillmann, 1994). The dimensions of concentric receptive fields are smallest in the retinal ganglion cells receiving their input from the central fovea, and they increase in size with increasing distance from the fovea. This is probably why the Hermann-Hering dots are not visible at the fixated intersection: both the center and the surround would fall between the squares and within the intersections, so that there would be no differential response from them and no illusory dots. If the separations are made sufficiently small, then the dots can be seen at the fixated intersections in Figure I.1-18. It is by this means that measuring the limiting dimensions of grids yielding the illusory dots has been used to estimate the sizes of "perceptive fields" at different eccentricities in human vision. However, doubt concerning this interpretation has been cast by the weaker or absent effect when curved rather than straight lines define the dark regions (Geier, Bernáth, Hudák, & Séra, 2008).

The grid illusion had, in fact, been reported before the publications of either Hermann or Hering. At a meeting of the British Association for the Advancement of Science, held at York in 1844, Brewster (1844a) reported a phenomenon communicated to him by a Rev. W. Selwyn:

> When a number of black parallel lines are intersected at right angles, so as to inclose a number of squares or rectangles, a white spot appears at the intersections of all the lines. In order to discover the cause of this phænomenon, Sir David Brewster made the experiment with the broad opake bars of an old-fashioned window opposed to the light of the sky. All along the bars he saw a whitish nebulous light, which was the complementary or accidental colour of the black bars seen simultaneously with the bars. The same luminosity was of course seen of equal intensity along all the bars, but at the crossings the intensity of its light was greatest, so as to produce the white spot already mentioned. Now this spot did not arise from any increased effect at the intersections, but from a *diminution* of the complementary luminosity at all other parts of the intersecting lines. This diminution of intensity arises from the action of the white squares or rectangles upon the retina tending to diminish the sensibility of that membrane along the parts corresponding to the black lines, and is always greatest in oblique vision. Its action is analogous to that which takes place when a strip of paper laid upon a green or any other coloured glass disappears when the eye is fixed upon a point an inch or two distant from the paper. Hence the luminous spots are brightest when not seen directly. (p. 8)

This is the interpretation given by Selwyn and reported by Brewster. It essentially relates the appearance of the illusory dots to simultaneous contrast and emphasizes the greater effect in peripheral vision, much as Hermann (1870) had done. (Perhaps we should now refer to the effect as the Selwyn grid!) This was a significant meeting of the British Association in visual terms because two descriptions of the fluttering hearts phenomenon were given in the same session of the optics section. Wheatstone (1844) had observed the shifting color effect in carpets, and Brewster (1844b) named it fluttering hearts. Wheatstone (1844)

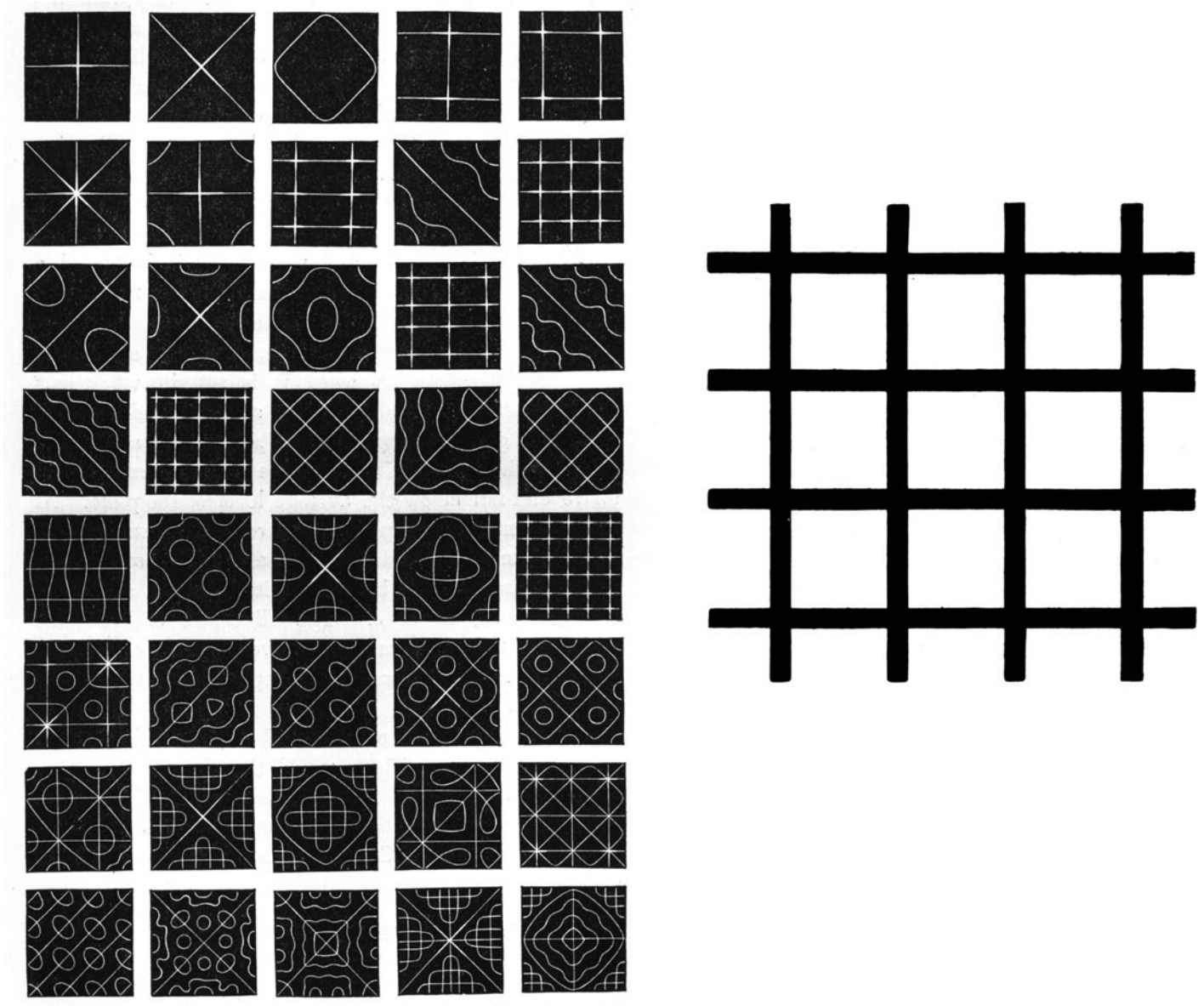


Figure I.1-18. Hermann and Hering grids. The figures in which (left) Hermann (1870) and (right) Hering (1907) saw the illusory dots at the intersections of the grids after which they are known. (Image © Nicholas Wade.)

noted that "the effects were more pronounced under low light conditions and he accounted for it by the eye retaining its sensibility for various colours during various lengths of time" (p. 288). That is, visual persistence was implicated in the phenomenon.

The combination and juxtapositioning of colors was known to be critical to make the metaphorical hearts flutter, although precise rules were not applied to them. Indeed, contrasting effects of brightness and color had long been known (Wade, 1998). Goethe (1810/1840) observed that: "A grey object on a black ground appears much brighter than the same object on a white ground. If both comparisons are seen together the spectator can hardly persuade himself that the two greys are identical" (p. 15). More precise observations were made by Michel Eugène Chevreul (1786–1889) in the context of fabric dyeing, because the same pigments were judged to appear quite different when presented together with other colored fabrics. This led Chevreul (1839) to formulate a law of simultaneous contrast: "When two contiguous colours are seen at the same time, they appear as dissimilar as possible, both with regard to their optical composition and their depth of tone" (Chevreul, 1861, p. 9). In addition, Chevreul made an arrangement of abutting grey rectangles, varying in density (Fig. I.1-19). He made the 10 distinct greys by painting the numbered sheets with the corresponding number of coats, so that area 1 had one coating and area 10 had 10. Not only did the greys contrast with their neighbors, but there were also light and dark bands on either side of the boundary; these bands were described in terms of the depth impressions they produced:

> In consequence of this contrast, the stripes seen from a proper distance, resemble channels rather than flat surfaces. For, in the stripes 2 and 3 for instance, the grey is

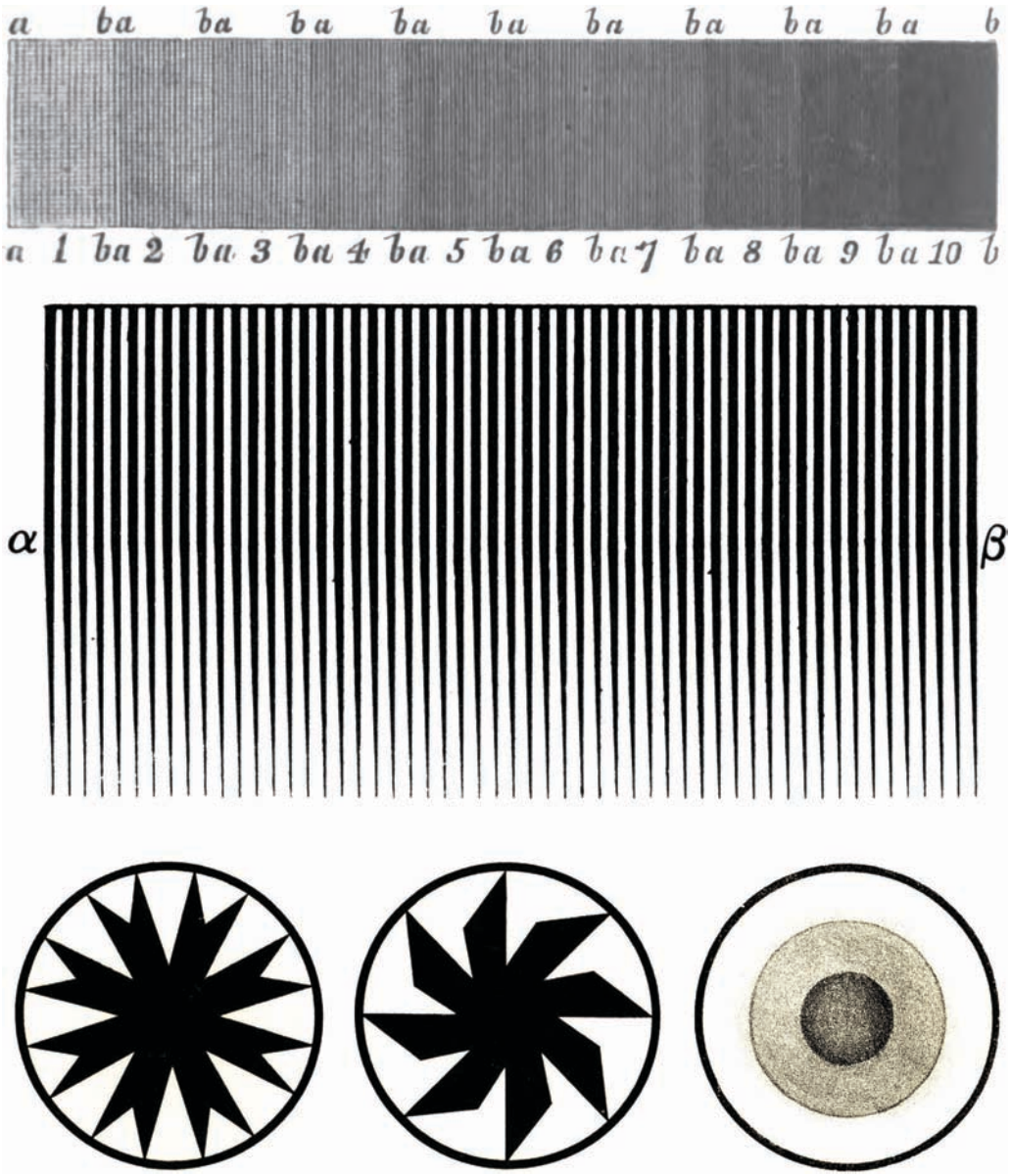


Figure I.1-19. Upper: Figure from Chevreul (1861) illustrating the arrangement for simultaneous contrast and border enhancement. There were 10 grey rectangles of differing tone: "Upon looking at the cardboard, it will be seen that the strips, instead of presenting flat tints, will each appear of a tone diminishing in intensity from the edge *a a* to the edge *b b*" (p. 6). Middle and lower figures are from Mach (1865). Similar effects can be seen in a stationary pattern consisting of ramps (middle). The apparent stripe in the figure "shows a dark line at *αβ*, and, from there downward the black lines thin gradually" (Ratliff, 1965, p. 260). Lower: Photographs were taken of the patterned discs when stationary and rotating, producing the brighter and darker rings at the border between the physically lighter and darker parts. (Image © Nicholas Wade.)

> weakened from the edge *aa* to the edge *bb*, presenting to the eye the same effect as if light fell upon a channelled surface. (p. 6)

Gradations of grey can be produced in several ways, and Mach (1865) employed this method as well as rotating luminance ramps (Fig. I.1-19). The effects can also be induced with static patterns in which there is a repetitive ramp varying in width, as Mach illustrated (Fig. I.1-19). The "channels" seen either side of grey boundaries are often referred to as Mach bands, together with the other methods of generating boundaries Mach produced (see Ratliff, 1965, for translations of Mach's six articles on border contrast). The border contrasts that Mach observed were described mathematically in terms of differential of luminance:

> The illumination of a retinal point will, in proportion to the difference between this illumination and the average of the illumination on neighboring points, appear brighter or darker, respectively, depending on whether the illumination of it is above or below that average. (Ratliff, 1965, pp. 321–322).

Ironically, when Mach (1906) looked back at his studies after 40 years, he ruefully commented that "My reports have remained unknown or unnoticed for a long time, such that, indeed, the facts concerned were again

discovered independently by observers more than 30 years later" (Ratliff, 1965, p. 322). Approximately the same time elapsed between Chevreul's (1839) description of border contrast and Mach's (1865) "rediscovery" of it. Geier and Hudák (2011) have examined characteristics seen in displays like those described by Chevreul.

ILLUSIONS IN THEORY

The new geometrical optical illusions were interpreted according to the visual theories of the day. However, prior to the explosion of experimental enquiries in the late 19th, a lucid classification and interpretation of illusions was

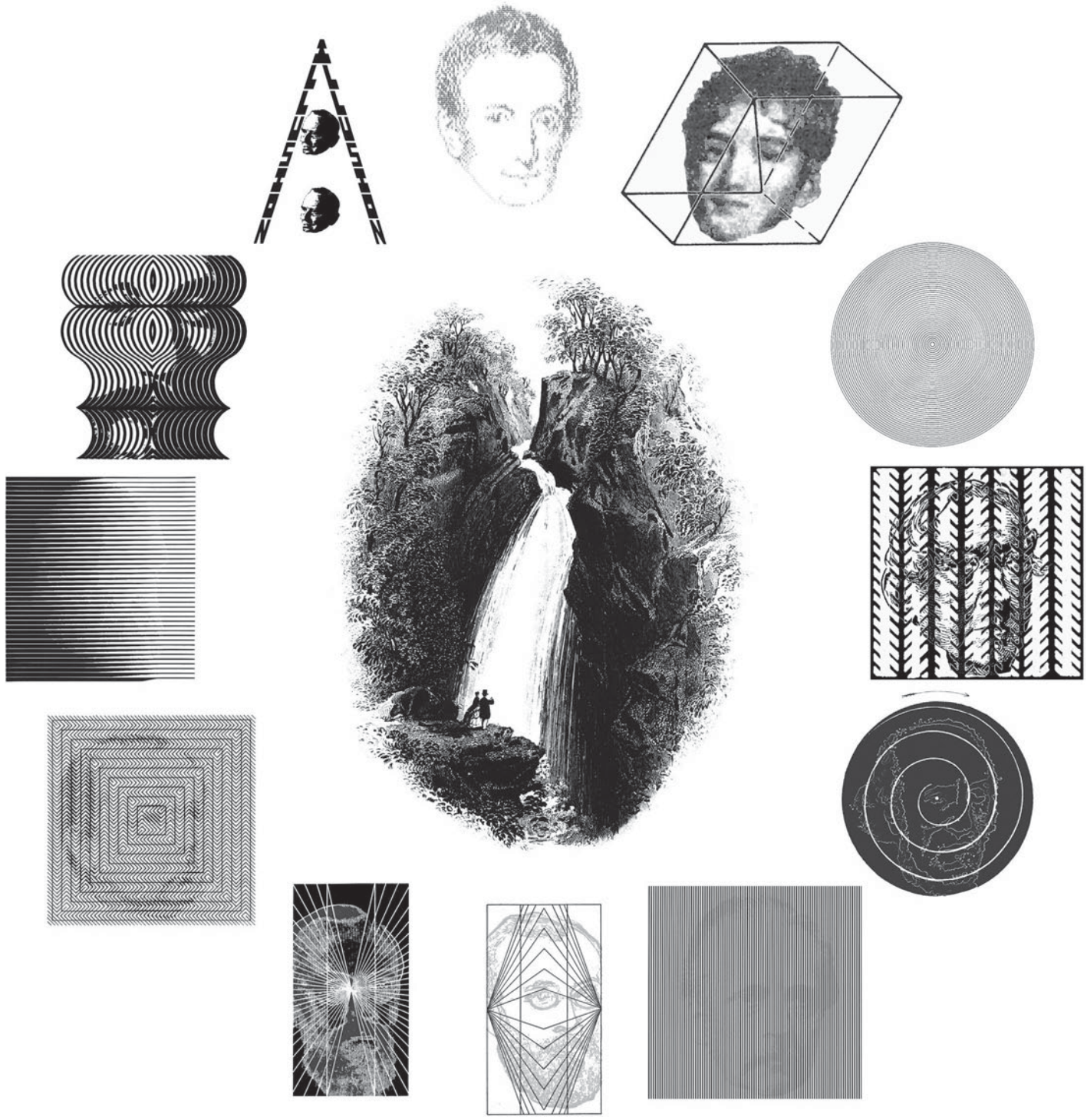


Figure I.1-20. *Nineteenth Century Pioneers of Visual Illusions* by Nicholas Wade. The central image is the Falls of Foyers where the waterfall illusion was observed. The portraits are (clockwise from the top in chronological sequence): Peter Mark Roget (1779–1869), Louis Albert Necker (1786–1861), Jan Evangelista Purkinje (1787–1869), Johann Christian Poggendorff (1796–1877), Joseph Antoine Ferdinand Plateau (1801–1883), Hermann Ludwig Ferdinand von Helmholtz (1821–894), Wilhelm Maximilian Wundt (1832–1920), Karl Ewald Konstantin Hering (1834–1918), Johann Karl Friedrich Zöllner (1834–1882), Ernst Mach (1838–1916), Ludimar Hermann (1838–1914), Franz Carl Müller-Lyer (1857–1916), and Mario Ponzo (1882–1960). (Image © Nicholas Wade.)

provided by Roget (1832, 1834). He considered that persisting images were visual illusions, and he applied his powers of classification to what he called such fallacious perceptions. His most detailed consideration of illusions was presented in his Gulstonian lectures on the laws of sensation and perception delivered in May, 1832, to the Royal College of Physicians. They were amplified in the second volume of his *Bridgewater Treatise* (Roget, 1834). Roget presented three classes of illusions—physical, physiological, and mental—together with examples of phenomena in them. "The first class included illusions in which an impression is really made on the organ of sense by an external cause; but in a way to which we have not been accustomed" (Roget, 1834, p. 532) to which mirages, mirror, and kaleidoscopic images belonged. "The second class comprehends those in which the cause of deception is more internal, and consists in the peculiar condition of the nervous surface receiving the impressions" (p. 535). Ocular spectra (afterimages) and persisting visual images were in this category. "The third class of fallacies comprehends those which are essentially mental in their origin, and are the consequences of errors in our reasoning powers" (p. 535). The three classes of illusions that Roget presented are similar to those that are applied in contemporary visual science of physical, physiological, and psychological (Gregory, 2009). Different approaches have been made to the interpretations of illusions in Roget's three categories. Illusions in the first class are generally interpreted in terms of geometrical optics. Interpretations of those in the second class have had some success with regard to the burgeoning knowledge of neural processing. Most attention is now directed at those of the third class, for which there are competing theories.

Since the 19th century, many attempts have been made to classify illusions in a manner that will facilitate interpreting them. Vicario (2011) listed 26 different classifications! Perhaps the simplest is the one proposed by Boring (1942): direction, orientation, and the rest. The common factors of the classifications are direction (orientation), size, contrast, assimilation, and perspective. Some systems are concerned with the distorted dimensions (like size and orientation), others involve possible underlying processes (like assimilation and contrast, eye movements), and yet others characterize the levels at which the illusions should be considered (like physical, physiological, and psychological).

There are many accounts of visual illusions from the 20th century (see Coren & Girgus, 1979; Day, 1972; Gregory, 2009; Luckiesh, 1922; Ninio, 1998; Robinson, 1972; Tolansky, 1964; Vicario, 2011; Wade, 1982, 2005a), not to mention the many entries in this *Compendium*. A more rounded view of contemporary research will be found in those sources than I can provide in this survey, which could be characterized as "visual illusions before geometrical optical illusions"! The final illustration (Fig. I.1-20) pays graphical homage to those 19th-century researchers who added to our understanding of visual illusions.

CONCLUSIONS

In order for an illusion to be so considered, two measurements are required. The most common are the physical characteristics of the stimulus and some suitable index of its perception. Illusions were studied in the late 19th century because they were not amenable to the extant physiology—hence their place in psychology. They also fostered the use of two-dimensional stimuli in perceptual experiments. This enchantment with pictorial images has been enhanced with the introduction of computer graphics: more complicated pictures can be produced. The seductive allure of the pictorial image has amplified the legacy of Kepler (1604): the belief that the first stage of vision is a two-dimensional image in the eye. Almost all the visual illusions now studied are presented on a two-dimensional screen, even when they involve motion. The question is then posed as to whether theories of visual illusions will exist independently of theories of vision that are concerned with three-dimensional space. Returning to Purkinje's (1823) statement: Will visual illusions reveal visual truths, or will visual truths reveal visual illusions?

REFERENCES

Adams, H. F. (1912). The autokinetic sensations. *Psychological Review Monograph Supplement*, *14*, 1–45.

Addams, R. (1834). An account of a peculiar optical phænomenon seen after having looked at a moving body. *London and Edinburgh Philosophical Magazine and Journal of Science*, *5*, 373–374.

Addams, R. (1835). Optische Täuschung nach Betrachtung eines in Bewegung begriffenen Körpers. *Annalen der Physik und Chemie*, *34*, 348.

Aitken, J. (1878). On a new variety of ocular spectrum. *Proceedings of the Royal Society of Edinburgh*, *10*, 40–44.

Alais, D., & Blake, R. (Eds.). (2005). *Binocular rivalry and perceptual ambiguity*. Cambridge, MA: MIT Press.

Aubert, H. (1865). *Physiologie der Netzhaut*. Breslau, Poland: Morgenstern.

Aubert, H. (1887): Die Bewegungsempfindungen. *Archiv für die gesamte Physiologie des Menschen und der Tiere*, *40*, 459–480.

Bacon, F. (1857). Sylva sylvarum: or a natural history. In J. Spedding, R. L. Ellis, & D. D. Heath (Eds.), *The works of Francis Bacon*, Vol. 2. London: Longman. (Original work published 1625)

Baltrusaitis, J. (1976). *Anamorphic art*. (W. J. Strachan, Trans.). London, Chadwyck-Healey.

Basler, A. (1910). Über das Sehen von bewegungen. V. Mitteilung. Untersuchungen über die simultane Scheinbewegung. *Archiv für die gesammte Physiologie des Menschen und der Thiere*, *132*, 131–142.

Beattie, W. (1838). *Caledonia illustrated in a series of views taken expressly for the work of W. H. Bartlett, T. Allom, and others*. London: Virtue.

Berkeley, G. (1709). *An essay towards a new theory of vision*. Dublin: Pepyat.

Bezold, W. von. (1884). Eine perspectivische Täuschung. *Annalen der Physik und Chemie*, *23*, 351–352.

Boring, E. G. (1930). A new ambiguous figure. *American Journal of Psychology*, *42*, 444–445.

Boring, E. G. (1942). *Sensation and perception in the history of experimental psychology*. New York: Appleton-Century.

Borschke, A., & Hescheles, L. (1902). Über Bewegungsnachbilder. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, *27*, 387–398.

Bowditch, H. P., & Hall, G. S. (1881). Optical illusions of motion. *Journal of Physiology*, *3*, 297–307.

Brandt, T. (2003). *Vertigo: Its multisensory syndromes* (2d ed.). London: Springer.

Breuer, J. (1874). Über die function der Bogengänge des Ohrlabyrinthes. *Wiener medicinische Jahrbuch*, *4*, 72–124.

Brewster, D. (1826). On the optical illusion of the conversion of cameos into intaglios, and of intaglios into cameos, with an account of other analogous phenomena. *Edinburgh Journal of Science*, *4*, 99–108.

Brewster, D. (1832). *Letters on natural magic addressed to Sir Walter Scott, Bart*. London: John Murray.

Brewster, D. (1844a). *A notice explaining the cause of an optical phenomenon observed by the Rev. W. Selwyn: Report of the British Association for the Advancement of Science.* Transactions Section 8. London: British Association for the Advancement of Science.

Brewster, D. (1844b). *On the same subject: Report of the British Association for the Advancement of Science.* Transactions Section 10. London: British Association for the Advancement of Science.

Brewster, D. (1845). *Notice of two new properties of the retina: Report of the British Association for the Advancement of Science.* Transactions Sections 9–10. London: British Association for the Advancement of Science

Brown, A. C. (1878). Cyon's researches on the ear. II. *Nature*, *18*, 657–659.

Brugger, P. (1999). One hundred years of an ambiguous figure: Happy birthday, duck/rabbit! *Perceptual & Motor Skills*, *89*, 973–977.

Burckhardt, T. (1976). *Art of Islam: Language and meaning.* (J. P. Hobson, Trans.). London: World of Islam Festival Trust.

Burton, H. E. (1945). The optics of Euclid. *Journal of the Optical Society of America*, *35*, 357–372.

Campbell, S. (1988). *The mosaics of Antioch*. Toronto: Pontifical Institute of Mediaeval Studies.

Carpenter, W. B. (1868). On the zoetrope and its antecedents. *The Student and Intellectual Observer*, *1*, 427–444.

Carr, H. A. (1910). The autokinetic sensation. *Psychological Review*, *17*, 42–75.

Charpentier, A. (1886). Sur une illusion visuelle. *Comptes Rendus*, *102*, 1155.

Chevreul, M. E. (1839). *De la loi du contraste simultané*. Paris: Pitois-Levrault.

Chevreul, M. E. (1861). *The laws of contrast of colour* (new ed.). (J. Spanton, Trans.). London: Routledge.

Coren, S., & Girgus, J. S. (1978). *Seeing is deceiving*. Hillsdale, NJ: Lawrence Erlbaum.

CW [Charles Wheatstone]. (1830). Contributions to the physiology of vision. No I. *Journal of the Royal Institution*, *1*, 101–117.

D'Arcy, Chevalier. (1765). Mémoire sur la durée de la sensation de la vue. *Mémoire de l'Académie des Science de Paris*, 439–451.

Darwin, E. (1794). *Zoonomia; or, the laws of organic life*. Vol. 1. London: Johnson.

Darwin, R. W. (1786). New experiments on the ocular spectra of light and colours. *Philosophical Transactions of the Royal Society*, *76*, 313–348.

Day, R. H. (1972). Visual spatial illusions: A general explanation. *Science*, *175*, 1335–1340.

Delboeuf, J. (1865a). Note sur certain illusions d'optique. *Bulletin de l'Academie Royal des Sciences, de Lettres et des Beaux-arts de Belgique*, *34*(19), 195–216.

Delboeuf, J. (1865b). Seconde note sur de nouvelles illusions d'optique. *Bulletin de l'Academie Royal des Sciences, de Lettres et des Beaux-arts de Belgique*, *34*(20), 70–97.

Delboeuf, J. (1892). Sur une nouvelles illusion d'optique. *Bulletin de l'Academie Royal des Sciences, de Lettres et des Beaux-arts de Belgique*, *24*, 545–558.

Dember, W. N. (1964). *Visual perception: The nineteenth century*. New York: Wiley.

Desaguliers, J. T. (1736). An explication of an experiment made in May 1735, as a farther confirmation of what was said in the paper given in January 30, 1734–5 to account for the appearance of the horizontal moon seeming larger than when higher. *Philosophical Transactions of the Royal Society, 39*, 392–394.

Descartes, R. (1902). *La dioptrique*. In C. Adam & P. Tannery (Eds.), *Oeuvres de Descartes* (Vol. 6, pp. 81–228). Paris: Cerf. (Original work published 1637).

Dewhurst, K. (1980). *Thomas Willis's Oxford lectures*. Oxford: Sandford.

Dewitz, B. von, & Nekes, W. (2002). *Ich sehe was, was du nicht siehst! Sehmachinen und Bilderwelten. Die Sammlung Werner Nekes*. Köln, Germany: Museum Ludwig.

Drake, S., & O'Malley, C. D. (1960). *The controversy on the comets of 1618*. Philadelphia: University of Pennsylvania Press.

Dunbabin, K. M. D. (2006). *Mosaics of the Greek and Roman world*. Cambridge, UK: Cambridge University Press.

Duncker, K (1929). Über induzierte Bewegung (Ein beitrag zur theorie optisch wahrgenommener Bewegung). *Psychologische Forschung*, *12*, 180–259.

Dvorak, V. (1870). Versuche über die Nachbilder von Reizveränderungen. *Sitzungsberichte der Wiener Akademie der Wissenschaften*, *61*, 257–262.

Edgerton, S. Y. (1975). *The Renaissance rediscovery of linear perspective*. New York: Basic Books.

Ehrenstein, W. (1925). Versuche über die Beziehungen zwischen Bewegung und Gestaltwahrnehmung. *Zeitschrift für Psychologie*, *96*, 305–352.

Ehrenstein, W. (1941). Über Abwandlungen der L. Hermannschen Helligkeitserscheinung. *Zeitschrift für Psychologie*, *150*, 83–91.

Euclid. (1895). *Optica*. (I. L. Heiberg, ed.). Leipzig, Germany: Teubner.

Exner, S. (1887). Einige Beobachtungen über Bewegungsnachbilde. *Centralblatt für Physiologie*, *1*, 135–140.

Exner, S. (1888). Über optische Bewegungsempfindungen. *Biologisches Centralblatt*, *8*, 437–448.

Exner, S. (1896). Über autokinetische Empfindungen. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, *12*, 313–330.

Faraday, M. (1831). On a peculiar class of optical deception. *Journal of the Royal Institution of Great Britain*, *1*, 205–223.

Fick, A. (1851). *Tractator de errore optico quodam asymetria bulbi oculi effecto*. Marburg, Germany: Koch.

Finger, S., & Wade, N. J. (2002). The neuroscience of Helmholtz and the theories of Johannes Müller: Part 2. Sensation and perception. *Journal of the History of the Neurosciences*, *11*, 234–254.

Fukuda, S. (2005). *Shigeo Fukuda masterworks*. Buffalo, NY: Firefly Books.

Galileo, G. (1606). *Considerazioni d'Alimberto Mauri sopra alcuni luoghi del discorso di Ludovico delle Colombe intorno alla stella apparita nel 1604.* Firenze, Italy: Caneo. (English translation: S. Drake in *Galileo against the philosophers in his Dialogue of Cecco di Ronchitti [1605] and Considerations of Alimberto Mauri [1606].* Los Angeles: Zeitlin & Ver Brugge, 1976).

Galileo, G. (1623). *Il Saggiatore, nel quale con bilancia esquisita e giusta*. Rome: Mascardi.

Geier, J., Bernáth, L., Hudák, M., & Séra, L. (2008). Straightness as the main factor of the Hermann grid illusion. *Perception*, *37*, 651–665.

Geier, J., & Hudák, M. (2011). Changing the Chevreul illusion by a background luminance ramp: Lateral inhibition fails at its traditional stronghold—A psychophysical refutation. *PLoS One*, 201, *6*(10), e26062.

Goethe, J. W. (1810). *Zur Farbenlehre*. Tübingen, Germany: Cotta.

Goethe, J. (1824). Das Sehen in subjectiven Hinsicht, von Purkinje, 1819. In J. Goethe, *Zur Naturwissenschaft überhaupt, besonders zur Morphologie* (pp. 102–117). Stuttgart, Germany: Cotta.

Gogel, W. C. (1977). Independent motion induced in separate portions of the visual field. *Bulletin of the Psychonomic Society*, *10*, 408–410.

Gregory, R. L. (2003). Delusions. *Perception*, *32*, 257–261.
Gregory, R. L. (2009). *Seeing through illusions.* Oxford: Oxford University Press.
Guilford, J. P., & Dallenbach, K. M. (1928). A study of the autokinetic sensation. *American Journal of Psychology*, *40*, 83–91.
Haller, A. von. (1786). *First lines of physiology*. Vol. 2. (W. Cullen, Trans.). Edinburgh, UK: Elliot.
Harris, J. (1775). *A treatise of optics: Containing elements of the science; in two books*. London: White.
Heller, D. (1988). On the history of eye movement recording. In G. Lüer, U. Lass, & J. Shallo-Hoffmann (Eds.), *Eye movement research: Physiological and psychological aspects* (pp. 37–51). Göttingen, Germany: Hogrefe.
Helmholtz, H. (1867). *Handbuch der physiologischen Optik.* In G. Karsten (Ed.), *Allgemeine Encyklopädie der Physik*. Vol. 9. Leipzig, Germany: Voss.
Helmholtz, H. von. (1896). *Handbuch der physiologischen Optik* (2d ed.). Hamburg: Voss.
Herschenson, M. (Ed.). (1989). *The moon illusion*. Hillsdale, NJ: Lawrence Erlbaum.
Hering, E. (1861). *Beiträge zur Physiologie. I. Vom Ortsinne der Nethaut*. Leipzig, Germany: Engelmann.
Hering, E. (1907). Vom simultanen Grenzkontrast. In T. Saemisch (Ed.), *Graefe-Saemisch Handbuch der gesamten Augenheilkunde* (2d ed., Vol. 3, Pt. 1, pp. 135–141). Leipzig, Germany: Engelmann.
Hermann, L. (1870). Eine Erscheinung des simultanen Contrastes. *Archiv für die gesammte Physiologie des Menschen und der Thiere*, *3*, 13–15.
Hill, W. E. (1915). My wife and my mother-in-law. *Puck*, *16*, 11.
Holt-Hansen, K. (1962). Oscillation experienced in perceived figures. *Historisk-filosofiske Meddelelser*, *39*, 1–48.
Horner, W. G. (1834). On the properties of the dædaleum, a new instrument of optical illusion. *London and Edinburgh Philosophical Magazine and Journal of Science*, *4*, 36–41.
Howard, I. P. (1982). *Human visual orientation*. New York: Wiley.
Humboldt, A. von. (1850). *Kosmos. Entwurf einer physischen Weltbeschreibung*. Vol. 3. Stuttgart, Germany: Cotta.
J.M. (1821). Account of an optical deception. *Quarterly Journal of Science, Literature, and Art*, *10*, 282–283.
Jastrow, J. (1899). The mind's eye. *Popular Science Monthly*, *54*, 301–312.
Jastrow, J. (1900). *Fact and fable in psychology*. Boston: Houghton Mifflin.
John, H. J. (1959). *Jan Evangelista Purkyne: Czech scientist and patriot, 1787–1869.* Philadelphia: American Philosophical Society.
Kanizsa, G. (1979). *Organization in vision: Essays on Gestalt perception*. New York: Praeger.
Keele, K. D. (1955). Leonardo da Vinci on vision. *Proceedings of the Royal Society of Medicine*, *48*, 384–390.
Kemp, M. (1990). *The science of art: Optical themes in Western art from Brunelleschi to Seurat*. New Haven, CT: Yale University Press.
Kepler, J. (1604). *Ad Vitellionem paralipomena*. Frankfurt: Marinium and Aubrii.
King-Hele, D. (1999). *Erasmus Darwin: A life of unequalled achievement*. London: De La Mer.
King-Hele, D. (Ed.). (2002). *Charles Darwin's The Life of Erasmus Darwin*. Cambridge, UK: Cambridge University Press.
Kleiner, A. (1878). Physiologisch-optische Beobachtungen. *Archiv für die gesammte Physiologie des Menschen und der Thiere*, *18*, 542–573.
Kondoleon, C. (1995). *Domestic and divine: Roman mosaics in the House of Dionysos*. Ithaca, NY: Cornell University Press.
Kruta, V. (1969). *J. E. Purkyně (1787–1869) physiologist: A short account of his contributions to the progress of physiology with a bibliography of his works*. Prague: Czechoslovak Academy of Sciences.
Kuhn, G., Amlani, A. A., & Rensink, R. A. (2008). Towards a science of magic. *Trends in Cognitive Sciences*, *12*, 349–354.
Kuhn, G., & Tatler, B. W. (2011). Misdirected by the gap: The relationship between inattentional blindness and attentional misdirection. *Consciousness and Cognition*, *20*, 432–436.
Kundt, A. (1863). Untersuchungen über Augenmass und optische Täuschungen. *Annalen der Physik und Chemie*, *120*, 118–158.
Lejeune, A. (Ed.). (1956). *L'Optique de Claude Ptolémée dans la version Latine d'après l'Arabe de l'Émir Eugène de Sicile*. Louvain: Université de Louvain.
Levi, D. M., & Schor, C. M. (1984). Spatial and velocity tuning of processes underlying induced motion. *Vision Research*, *24*, 1189–1196.
Levy, J. (1972). Autokinetic illusion: A systematic review of theories, measures and independent variables. *Psychological Bulletin*, *78*, 457–474.
Lingelbach, B. (2008). Der Necker-Würfel—oder was ist vorne. *Klinische Monatsblätter für Augenheilkunde*, *225*, 253–254.
Lingelbach, B. (2012a). Illusionäre Konturen. *Focus*, *31*(1), 44–46.
Lingelbach, B. (2012b). Sant' Ignazio: Himmlische perspective. *Focus*, *31*(12), 28–30.
Lingelbach, B., & Ehrenstein, W. H. (2002). Das Hermanngitter und die Folgen. *Deutsche Optikerzeitung*, *5*, 14–20.
Lingelbach, B., & Wade, N. (2010). Penrose "darts" and "kites" on medieval buildings. *Perception*, *39*, 115.
Luckiesh, M. (1922). *Visual illusions*. London: Constable.
Lucretius. (1975). *De rerum natura*. (W. H. D. Rouse, Trans.). Cambridge, MA: Harvard University Press.
MacCurdy, E. (1938). *The notebooks of Leonardo da Vinci*. Vol. 1. London: Cape.
Mach, E. (1865). Ueber die physiologische Wirkung raümlichen Vertheilung des Lichtreize auf die Netzhaut. *Sitzungsberichte der mathematisch-naturwissenschaftlichen Classe der kaiserlichen Akademie der Wissenschaften*, *52*, 303–322.
Mach, E. (1866). Ueber die physiologische Wirkung raümlich vertheilter Lichtreize. (Dritte Abhandlung). *Sitzungsberichte der mathematisch-naturwissenschaftlichen Classe der kaiserlichen Akademie der Wissenschaften*, *54*, 393–408.
Mach, E. (1875). *Grundlinien der Lehre von den Bewegungsempfindungen*. Leipzig, Germany: Engelmann.
Mach, E. (1906). Ueber den Einfluss raümlich und zeitlich variierender Lichtreize auf die Gesichtswahrnehmung. *Sitzungsberichte der mathematisch-naturwissenschaftlichen Classe der kaiserlichen Akademie der Wissenschaften*, *115*, 633–648.
Mack, A. (1986). Perceptual aspects of motion in the frontal plane. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Vol 1. Sensory processes and perception* (pp. 17–38). New York: Wiley.
Mannoni, L. Nekes, W., & Warner, M. (Eds.). (2004). *Eyes, lies and illusions*. London: Hayward Gallery.
Mather, G., Verstraten, F., & Anstis, S. (Eds.). (1998). *The motion after-effect: A modern perspective*. Cambridge, MA: MIT Press.
Martinez-Conde, S., & Macknik, S. L. (2008). Magic and the brain. *Scientific American*, *299*(6), 72–79.
Matousek, O. (1961). J. E. Purkynes Leben und Tätigkeit im Lichte der Berliner und Prager Archive. *Nova Acta Leopoldina*, *24*, 109–129.
Molyneux, W. (1687). Concerning the apparent magnitude of the sun and moon, or the apparent distance of two stars, when nigh the horizon, and when higher elevated. *Philosophical Transactions of the Royal Society*, *16*, 314–323.
Müller, J. (1826a). *Über die phantastischen Gesichtserscheinungen*. Coblenz, Germany: Hölscher.
Müller, J. (1826b). *Zur vergleichenden Physiologie des Gesichtssinnes des Menschen und der Thiere, nebst einen Versuch über die Bewegung der Augen und über den menschlichen Blick*. Leipzig, Germany: Cnobloch.

Müller, J. (1838). *Handbuch der Physiologie des Menschen*. Vol. 2. Coblenz, Germany: Hölscher.
Müller, J. (1843). *Elements of physiology* (2d ed.). Vol. 2. (W. Baly, Trans.). London: Taylor and Walton.
Müller, J. (2003). *Müller's Elements of physiology.*4 vols. (W. Baly, Trans.). Bristol, UK: Thoemmes.
Müller-Lyer, F. C. (1889). Optische Urtheilstäuschungen. *Archiv fur Anatomie und Physiologie, Physiologische Abteilung, 2*, 263–270.
Necker, L. A. (1832). Observations on some remarkable phenomena seen in Switzerland; and an optical phenomenon which occurs on viewing a figure of a crystal or geometrical solid. London and Edinburgh. *Philosophical Magazine and Journal of Science*, *1*, 329–337.
Newton, I. (1717). *Opticks: or, a treatise of the reflections, refractions, inflections and colours of light* (2d ed.). London: Innys.
Nicolas, S. (1995). Joseph Delboeuf on visual illusions: A historical sketch. *American Journal of Psychology*, *108*, 563–574.
Ninio, J. (1998). *The science of illusions*. (F. Philip, Trans.). Ithaca, NY: Cornell University Press.
Oppel, J. J. (1855). Über geometrisch-optische Täuschungen. *Jahresbericht des physikalischen Vereins zu Frankfurt am Main*, *1854–1855*, 37–47.
Oppel, J. J. (1856). Neue Beobachtungen und Versuche über eine eigentümliche, noch wenig bekannte Reaktionsthätigkeit des menschlichen Auges. *Annalen der Physik und Chemie*, *99*, 540–561.
Oppel, J. J. (1861). Ueber geometrisch-optische Täuschungen (Zweite Nachlese). *Jahresberichte des physikalischen Vereins zu Frankfurt am Main, 1860–1861*, 26–37.
Papathomas, T. V. (2007). Art pieces that "move" in our minds—An explanation of illusory motion based on depth reversal. *Spatial Vision*, *21*, 79–95.
Paris, J. A. (1827). *Philosophy in sport made science in earnest! Being an attempt to illustrate the first principles of natural philosophy by the aid of popular toys and sports.* Vol.3. London: Longman.
Petry, S., & Meyer, G. E. (1987). *The perception of illusory contours*. New York: Springer.
Phillips, D., & Wade, N. J. (2014). The elusive Johann Joseph Oppel (1815–1894). *Perception*, *43*, 869–872.
Piccolino, M., & Wade, N.J. (2006a). Flagging early examples of ambiguity. I. *Perception*, *35*, 861–864.
Piccolino, M., & Wade, N. J. (2006b). Flagging early examples of ambiguity. II. *Perception*, *35*, 1003–1006.
Piccolino, M., & Wade, N.J. (2007). *Insegne ambiguë. Percorsi obliqui tra storia, scienza e arte, da Galileo a Magritte.* Pisa, Italy: Edizioni ETS.
Piccolino, M., & Wade, N. J. (2008). Galileo's eye: A new vision of the senses in the work of Galileo Galilei. *Perception*, *37*, 1312–1340.
Piccolino, M., & Wade, N.J. (2014). *Galileo's visions: Piercing the spheres of the heavens by eye and mind.* Oxford: Oxford University Press.
Pinna, B., & Reeves, A. (2009). From perception to art: How vision creates meanings. *Spatial Vision*, *22*, 225–272.
Pind, J. L. (2014). *Edgar Rubin and psychology in Denmark*. Heidelberg, Germany: Springer.
Pirenne, M. (1970). *Optics, painting and photography*. Cambridge, UK: Cambridge University Press.
Plateau, J. (1829). *Dissertation sur quelques propriétés des impressions produites par la lumière sur l'organe de la vue*. Liége, Belgium: Dessain.
Plateau, J. (1833). Des illusions sur lesquelles se fonde le petit appareil appelé récemment Phénakisticope. *Annales de Chimie et de Physique de Paris*, *53*, 304–308.
Plateau, J. (1836). Notice sur l'anorthoscope. *Bulletins de l'Académie Royale des Sciences et Belles-Lettres de Bruxelles*, *3*, 7–10.
Plateau, J. (1849). Quatrième note sur de nouvelles applications curieurses de la persistence des impressions de la rétine. *Bulletins de l'Académie Royale des Sciences, des Lettres et des Beaux-Arts de Bruxelles*, *16*, 254–260.
Plateau, J. (1850). Vierte Notiz über neue, sonderbare Anwendungen des Verweilens der Eindrücke auf die Netzhaut. *Annalen der Physik und Chemie*, *80*, 287–292.
Plateau, J. (1878). Bibliographie analytique des principaux phénomènes de la vision, depuis les temps anciens jusqu'a la fin du XVIIIe siècle. Première section. Persistance des impressions sur la rétine. *Mémoires de l'Académie Royale de Belgique*, *42*, 1–59.
Plug, C., & Ross, H. (1989). Historical review. In M. Herschenson (Ed.), *The moon illusion* (pp. 5–27). Hillsdale, NJ: Lawrence Erlbaum.
Ponzo, M. (1912). Rapports entre quelques illusions visuelle de contrast angulaire et l'appréciation de grandeur des astres à l'horizon. *Archives Italiennes de Biologie*, *58*, 327–329.
Porterfield, W. (1759). *A treatise on the eye, the manner and phœnomena of vision*. Edinburgh, UK: Hamilton and Balfour.
Post, R. B., & Leibowitz, H. W. (1985). A revised analysis of the role of efference in motion perception. *Perception*, *14*, 631–643.
Purghé, F., & Coren, S. (1992). Subjective contours 1900–1990: Research trends and bibliography. *Perception & Psychophysics*, *51*, 291–304.
Purkinje, J. (1820). Beyträge zur näheren Kenntniss des Schwindels aus heautognostischen Daten. *Medicinische Jahrbücher des kaiserlich-königlichen öesterreichischen Staates*, *6*, 79–125.
Purkinje, J. (1823). *Beobachtungen und Versuche zur Physiologie der Sinne. Beiträge zur Kenntniss des Sehens in subjectiver Hinsicht*. Prague: Calve.
Purkinje, J. (1825). *Beobachtungen und Versuche zur Physiologie der Sinne. Neue Beiträge zur Kenntniss des Sehens in subjectiver Hinsicht*. Berlin: Reimer.
Raetz, M. (2002). *Nothing is lighter than light*. Paris: Maison Européenne de la Photographie.
Ratliff, F. (1965). *Mach bands: Quantitative studies on neural networks in the retina*. San Francisco, CA: Holden-Day.
Reusner, N. (1589). *Icones sives imagines vivae, literis*. Basel, Germany: Waldkirch.
Rittenhouse, D. (1786). Explanation of an optical deception. *Transactions of the American Philosophical Society*, *2*, 37–42.
Robinson, J. O. (1972). *The psychology of visual illusion*. London: Hutchinson.
Roget, P. M. (1825). Explanation of an optical deception in the appearance of the spokes of a wheel seen through vertical apertures. *Philosophical Transactions of the Royal Society*, *115*, 131–140.
Roget, P. M. (1832). Abstract of the Gulstonian lectures. *London Medical Gazette*, *10*, 273–282.
Roget, P. M. (1834). *Animal and vegetable physiology considered with reference to natural theology: Bridgewater treatise V*. Vol. 2. London: Pickering.
Ross, H. E., & Plug, C. (2002). *The mystery of the moon illusion.* Oxford: Oxford University Press.
Ross, H. E., & Ross, G. W. (1976). Did Ptolemy understand the moon illusion? *Perception*, *5*, 377–385.
Ross, W. D. (Ed.). (1931). *The works of Aristotle*. Vol. 3. Oxford: Clarendon.
Rubin. E. (1915). *Synsoplevede figurer*. Copenhagen: Gyldendalske.
Rubin, E. (1921). *Visuell wahrgenommene Figuren*. Copenhagen: Gyldendalske.
Rubin, E. (1958). Figure and ground. In D. C. Beardslee & M. Wertheimer (Eds.), *Readings in perception* (pp. 194–203). New York: van Nostrand.
Sabra, A. I. (1987). Psychology versus mathematics: Ptolemy and Alhazen on the moon illusion. In E. Grant & J. E. Murdoch (Eds.), *Mathematics and its application to natural philosophy*

in the Middle Ages (pp. 217–247). Cambridge, UK: Cambridge University Press.

Sabra, A. I. (Ed. and Trans.). (1989). *The Optics of Ibn Al-Haytham. Books I–III. On direct vision*. London: Warburg Institute.

Scheiner, C. (1619). *Oculus, hoc est fundamentum opticum*. Innsbruck, Austria: Agricola.

Schiller, P. H., & Carvey, C. E. (2005). The Hermann grid illusion revisited. *Perception*, *34*, 1375–1397.

Schröder, H. (1858). Ueber eine optische Inversion bei Betrachtung verkehrter, durch optische Vorrichtung entworfener, physicher Bilder. *Annalen der Physick und Chemie*, *181*, 298–311.

Schumann, F. (1900). Beiträge zur Analyse der Gesichtswahrnehmungen. Erste Abteilung. Einige Beobachtungen über die Zusammenfassung von Gesichtseindrucken zu Einheiten. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, *23*, 1–32.

Schweizer, G. (1857). Über das Sternschwanken. I. *Bulletin de la Société Impériale des Naturalistes de Moscou, 30*, 440–457.

Schweizer, G. (1858). Über das Sternschwanken. II. *Bulletin de la Société Impériale des Naturalistes de Moscou*, *31*, 477–500.

Skolnick, A (1940). The role of eye movements in the autokinetic phenomenon. *Journal of Experimental Psychology*, *26*, 373–393.

Smith, A. M. (1996). *Ptolemy's theory of visual perception: An English translation of the* Optics *with introduction and commentary*. Philadelphia: American Philosophical Society.

Smith, R. (1738). *A compleat system of opticks in four books*. Cambridge, UK: Author.

Smith, R. (2004). *A compleat system of opticks in four books*. Bristol, UK: Thoemmes Continuum.

Spillmann, L. (1994). The Hermann grid illusion: A tool for studying human perceptive field organization. *Perception*, *23*, 691–708.

Stampfer, S. (1833). *Die stroboskopischen Scheiben oder optische Zauberscheiben, deren Theorie und Wissenschaftliche Anwendung*. Vienna: Trentsensky & Vieweg.

Strong, D. S. (1979). *Leonardo on the eye*. New York: Garland.

Swanston, M. T., & Wade, N. J. (1992). Motion over the retina and the motion aftereffect. *Perception*, *21*, 569–582.

Swanston, M. T., & Wade, N. J. (1994). A peculiar optical phænomenon. *Perception*, *23*, 1107–1110.

Szily, A. von. (1905). Bewegungsnachbild und Bewegungskontrast. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, *38*, 81–154.

Tatler, B. W, & Wade, N. J. (2003). On nystagmus, saccades, and fixations, *Perception*, *32*, 167–184.

Taws, R. (2007). Trompe l'oeil and trauma: Money and memory after the Terror. *Oxford Art Journal, 30*, 355–376.

Thiéry, A. (1895a). Ueber geometrisch-optische Täuschungen. *Philosophische Studien*, *11*, 307–370.

Thiéry, A. (1895b). Ueber geometrisch-optische Täuschungen. *Philosophische Studien*, *11*, 603–620.

Thiéry, A. (1896). Ueber geometrisch-optische Täuschungen. *Philosophische Studien*, *12*, 67–125.

Thompson, S. P. (1877). *Some new optical illusions: Report of the British Association for the Advancement of Science.* Transactions Section 32. London: British Association for the Advancement of Science.

Thompson, S. P. (1880). Optical illusions of motion. *Brain 3*, 289–298.

Titchener, E. B. (1901). *Experimental psychology: A manual of laboratory practice.* Vol. 1. *Qualitative experiments*. Part 1. Student's Manual. New York: Macmillan.

Tolansky, S. (1964). *Optical illusions*. London: Pergamon Press.

Tyndall, J. (1869). *Der Schall: Acht Vorlesungen Gehalten in der Royal Institution von Gross-Britannien*. Braunschweig, Germany: Vieweg

Verriest, G. (1990). Life, eye disease and work of Joseph Plateau. *Documenta Ophthalmologica*, *74*, 9–20.

Vicario, G. B. (2008). Optical-geometrical illusions: The nomenclature. *Gestalt Theory*, *30*, 168–180.

Vicario, G. B. (2011). *Illusioni ottico-geometriche. Una rassegna di problemi*. Venice: Istituto Veneto di Scienze, Lettere ed Arte.

Wackermann, J. (2011). Filled space expansion: Constants, variants and determinants of the Oppel-Kundt phenomenon. In D. Algom, D. Zakay, E. Chajut, S. Shaki, Y. Mama, & V. Shakuf (Eds.). *Fechner Day 2011* (pp. 47–52). Ra'anana, Israel: International Society for Psychophysics.

Wade, N. (1982). *The art and science of visual illusions*. London: Routledge and Kegan Paul.

Wade, N. (1990). *Visual allusions: Pictures of perception*. Hove, UK: Lawrence Erlbaum.

Wade, N. J. (1994). A selective history of the study of visual motion aftereffects. *Perception*, *23*, 1111–1134.

Wade, N. J. (1998). *A natural history of vision*. Cambridge, MA: MIT Press.

Wade, N. J. (2000) William Charles Wells (1757–1817) and vestibular research before Purkinje and Flourens. *Journal of Vestibular Research*, *10*, 127–137.

Wade, N. J. (2002). Erasmus Darwin (1731–1802). *Perception*, *31*, 643-650.

Wade, N. J. (2003). *Destined for distinguished oblivion: The scientific vision of William Charles Wells (1757–1817).* New York: Kluwer/Plenum.

Wade, N. J. (2004). Philosophical instruments and toys: Optical devices extending the art of seeing. *Journal of the History of the Neurosciences*, *13*, 102–124.

Wade, N. J. (2005a). *Perception and illusion: Historical perspectives*. New York: Springer.

Wade, N. J. (2005b). Sound and sight: Acoustic figures and visual phenomena. *Perception*, *34*, 1275–1290.

Wade, N. J. (2007a). Artful visions. *Spatial Vision*, *21*, 27–53.

Wade, N. J. (2007b). Galileo and the senses: Vision and the art of deception. *Galilæana*, *4*, 259–288.

Wade, N. J. (2007c). The vision of William Porterfield. In H. Whitaker, C. U. M. Smith, & S. Finger (Eds.), *Brain, mind, and medicine: Essays in 18th century neuroscience* (pp. 163–176). New York: Springer.

Wade, N. J. (2010a). Vision and the word. *Perception*, *39*, 1157–1162.

Wade, N. J. (2010b). Pioneers of eye movement research. *i-Perception*, *1*(2), 33–68.

Wade, N. J. (2010c). The Darwins and Wells: From revolution to evolution. *Journal of the History of the Neurosciences*, *19*, 85–104.

Wade, N. J. (2011). Neurognostics answer: Science and sense. *Journal of the History of the Neurosciences*, *20*, 74–79.

Wade, N. J. (2012a). Artistic precursors of Gestalt principles. *Gestalt Theory*, *34*, 329–348.

Wade, N. J. (2012b). Wheatstone and the origins of moving stereoscopic images. *Perception*, *41*, 901–924.

Wade, N. J. (2014). Geometrical optical illusionists. *Perception*, *43*, 846–868.

Wade, N. J., & Brožek, J. (2001). *Purkinje's vision: The dawning of neuroscience*. Mahwah, NJ: Lawrence Erlbaum.

Wade, N. J., & Heller, D. (1997). Scopes of perception: The experimental manipulation of space and time. *Psychological Research*, *60*, 227–237.

Wade, N. J., & Heller, D. (2003). Visual motion illusions, eye movements, and the search for objectivity. *Journal of the History of the Neurosciences*, *12*, 376–395.

Wade, N. J., & Hughes, P. (1999). Fooling the eyes: *Trompe l'oeil* and reverse perspective. *Perception*, *28*, 1115–1119.

Wade, N. J., & Nekes, W. (2005). The two faces of Rex Whistler (1905–1944). *Perception*, *34*, 639–644.
Wade, N. J., & Piccolino, M. (2009). Seeing stars. *Perception*, *38*, 1581–1584.
Wade, N. J., Campbell, R. N., Ross, H. E., & Lingelbach, B. (2010). Necker in Scotch Perspective. *Perception*, *39*, 1–4.
Wade, N. J., & Swanston, M. (2013). *Visual perception: An introduction* (3d ed.). Hove, UK: Psychology Press.
Wade, N. J., Swanston, M. T., Howard, I. P., Ono, H., & Shen, X. (1991). Induced rotary motion and ocular torsion. *Vision Research*, *31*, 1979–1983.
Wade, N. J., &Tatler, B. W. (2005). *The moving tablet of the eye: The origins of modern eye movement research*. Oxford: Oxford University Press.
Wade, N. J., & Tatler, B. W. (2011). Origins and applications of eye movement research. In S. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *Oxford handbook on eye Movements* (pp. 17–46). Oxford: Oxford University Press.
Wade, N. J., & Verstraten, F. A. J. (1998). Introduction and historical overview. In G. Mather, F. Verstraten, & S. Anstis (Eds.), *The motion after-effect: A modern perspective* (pp. 1–23). Cambridge, MA: MIT Press.
Wallin, J. E. W. (1905). *Optical illusions of reversible perspective: A volume of historical and experimental researches*. Princeton, NJ: Author.
Wallis, J. (1687). The sentiments of the Reverend and Learned Dr. Wallis R. S. Soc. upon the aforesaid appearance. *Philosophical Transactions of the Royal Society*, *16*, 323–329.
Wells, W. C. (1792). *An essay upon single vision with two eyes: Together with experiments and observations on several other subjects in optics*. London: Cadell.
Wells, W. C. (1794a). Reply to Dr. Darwin on vision. *The Gentleman's Magazine and Historical Chronicle*, *64*, 794–797.
Wells, W. C. (1794b). Reply to Dr. Darwin on vision. *The Gentleman's Magazine and Historical Chronicle*, *64*, 905–907.
Wheatstone, C. (1827). Description of the kaleidophone, or phonic kaleidoscope: A new philosophical toy, for the illustration of several interesting and amusing acoustical and optical phenomena. *Quarterly Journal of Science, Literature and Art*, *23*, 344–351.
Wheatstone, C. (1838). Contributions to the physiology of vision—Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society*, *128*, 371–394.
Wheatstone, C. (1844). *On the singular effect of the juxtaposition of certain colours under particular circumstances: Report of the British Association for the Advancement of Science*. Transactions Section 10. London: British Association for the Advancement of Science.
Willis, T. (1672). *De anima brutorum*. London: Wells & Scott.
Wohlgemuth, A. (1911). On the after-effect of seen movement. *British Journal of Psychology. Monograph Supplement*, *1*, 1–117.
Wright, E. (1992). The original of E. G. Boring's young girl/mother-in-law drawing and its relation to the pattern of a joke. *Perception*, *21*, 273–275.
Wundt, W. (1874). *Grundzüge der physiologischen Psychologie*. Leipzig, Germany: Engelmann.
Wundt, W. (1898). *Die geometrisch-optischen Täuschungen*. Leipzig, Germany: Teubner.
Young, L. R., Henn, V., & Scherberger, H. (2001). *Fundamentals of the theory of movement perception by Dr. Ernst Mach*. New York: Kluwer/Plenum.
Young, T. (1800). Outlines of experiments and enquiries respecting sound and light. *Philosophical Transactions of the Royal Society*, *90*, 106–150.
Young, T. (1807). *A course of lectures on natural philosophy and the mechanical arts*. London: Johnson.
Young, T. (2002). *Thomas Young's lectures on natural philosophy and the mechanical arts*. Bristol, UK: Thoemmes.
Zöllner, F. (1860). Über eine neue Art von Pseudoskopie und ihre Beziehungen zu den von Plateau und Oppel beschreibenen Bewegungsphänomenen. *Annalen der Physik und Chemie*, *110*, 500–523.
Zöllner, F. (1862). Ueber eine neue Art anorthoscopische Zerrbilder. *Annalen der Physik und Chemie*, *117*, 477–484.

Chapter 2
Cross-Cultural Studies of Illusions

J. B. Deręgowski

INTRODUCTION

It is conceptually easy to distinguish three grand vectors that jointly determine an observer's visual perception: (a) culture: manmade entities, social usages such as language and manners, tools, and ways of dealing with the environment; (b) environment: the very surroundings in which the observer lives, the surrounding landscape, its elevation above sea level, the opportunities it offers, and the limits it imposes; and (c) the genetic makeup of the observer.

Cross-cultural psychology attempts to investigate the influence of these vectors, but because these vectors are not mutually independent, empirical investigation of their individual influence is practically impossible. Reports of cross-cultural studies of perception of illusions do not, therefore, as the following review will show, present a clear answer as to the true significance of each of these vectors but offer at best an indication of their relative importance.

Cross-cultural studies are of interest primarily because they demonstrate the extent of perceptual variation which would not be observable if the samples were derived from a single population. They are also of interest because they show how variation of various factors affects perception.

The studies of visual illusions which lie at the very beginnings of cross-cultural psychology and which, by collecting data outside laboratories and by comparing populations; not merely assessing performance of individuals, pioneered a new approach to studies of perception in general are those of the Cambridge Anthropological Expedition to Torres Straits. A. C. Haddon, who led the expedition, felt (a feeling then rare among anthropologists) that "no investigation of a people was complete that did not embrace a study of their psychology" (Haddon, 1901, preface p. v) and invited W. H. R. Rivers to carry out an empirical investigation of psychological and physiological characteristics of the population of the Straits. Rivers was a talented medical man who in the course of his life made significant contributions to medicine, psychology, and anthropology. His psychological investigations in Torres Straits were concerned with vision and embraced, medical examinations apart, both colour perception and, under the heading "visual spatial perception," studies of binocular vision, bisection and trisection of lines, and visual illusions (Rivers, 1901, pp. 97–132).

Zusne (1968) presents a summary of studies of visual illusions from the first studies until 1967. The index used by him is the number of relevant publications. This shows that up to the time of the sailing of the Cambridge Expedition from London (March, 1898) the proportion of studies concerned with illusions was consistently larger than that of other studies of visual perception, and larger than it ever became in the future, and that the number of studies of illusions was steadily growing. Illusions were clearly thought to be of particular importance to understanding visual perception, and this probably persuaded Rivers to study them. There was also another, philosophical influence deriving from the Spencerian assertion that dominance of lower intellectual life hinders higher intellectual life—an assertion which provides the reason for concentrating on sensory processes within cognition. Herbert Spencer (1820–1903) was the great sociological thinker of the second half of the 19th century, and Rivers was initially much affected by his ideas. This effect decreased as empirical evidence gathered from remote populations showed them to be questionable (Jahoda, 1992).

Rivers in Torres Straits probed the illusions shown in Figure I.2-1, namely, (a) horizontal–vertical, (b) Müller-Lyer, (c) Zöllner, (d) Hering, (e) Wundt, (f) Jastrow, and (g) Bowdich and Hall. Quantitative measures of responses to the first two were obtained, while responses to the others were observed and assessed qualitatively. The two illusions chosen for thorough investigation were, as will be seen, those destined in the future to be at the centre of cross-cultural studies. In addition to the illusion stimuli, Rivers also used a pictorial stimulus (which he termed an illusion). This was a picture of a box bridge with two persons standing on it. The bridge was drawn in converging perspective so that together with the two standing figures the picture presented a stimulus similar to the Ponzo illusion. Responses to this stimulus were, however, meager, possibly, Rivers thought, because the two silhouettes depicted on the bridge, a top-hatted gentleman in a long winter coat and a schoolboy with a satchel and wearing a traditional English school cap, were entirely unfamiliar to the islanders. Apart from its obvious kinship to the Ponzo illusion, a stimulus investigated intensely later, the use of the box bridge stimulus also anticipated cross-cultural investigations of the problems of pictorial perception.

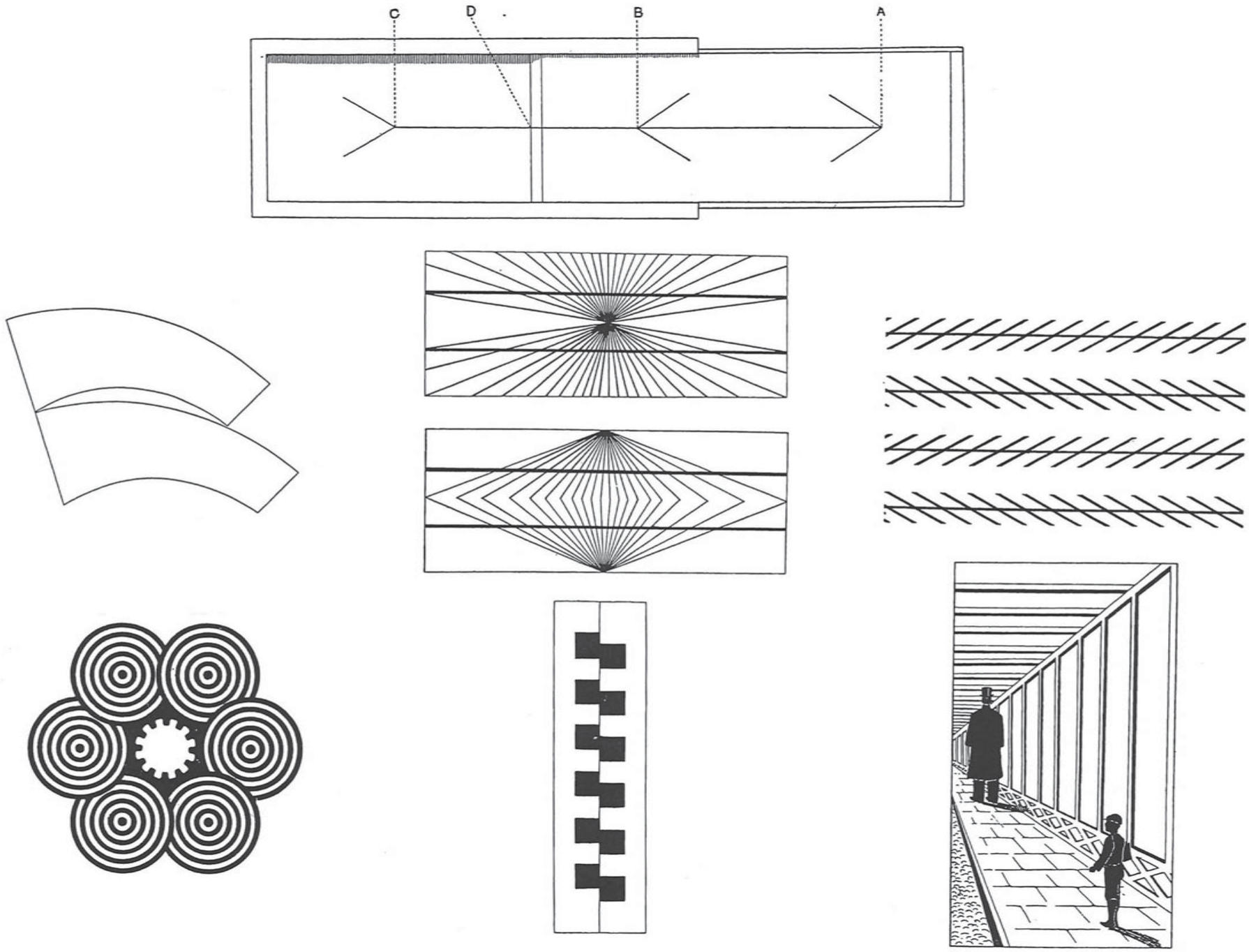


Figure I.2-1. Stimuli used by Rivers (1901). The slide shown at the top is of the Torres Straits version (A–B and B–C are the compared lengths; D is the point of overlap). The horizontal–vertical illusion (not shown) was tested by requesting subjects to draw a line equal to a stimulus line, either at the end, at the middle, or through the middle of the stimulus line, and later in India by means of an apparatus involving adjustable lines. (Other figures shown were not used to obtain quantitative data but in an exploratory manner.) Adapted from Rivers (1901).

Subsequently, Rivers extended his studies of illusions to three other exotic populations: the Uralis, the Sholagas, and the Todas of the Indian subcontinent. He did not employ the same methods in all of his studies but rather modified them to make the findings more reliable. These changes do not, however, affect the overall validity of the results.

The findings from Torres Straits are presented here together with findings from Southern India; Uralis and Sholagas, who were "people wilder than any previously examined" (Rivers, 1903); Todas; and an English sample (Rivers, 1905).

The results obtained by Rivers led him to the following conclusions. All populations tested were prone to illusions but differed in their sensitivity, and these differences varied with illusion. Thus the horizontal–vertical illusion was found to be distinct more marked to the Toda and the Papuans than to the Europeans, and European responses were less consistent (an observation confirmed by a later study by Hicks and Rivers [1908]). In contrast, the Müller-Lyer illusion was found to be less marked to the "less civilized," and this was more distinctly in the case of the Papuans than the Toda. Furthermore, present-day reanalysis of the data obtained from "the people wilder than any" made possible by the advance of statistical techniques confirms Rivers's speculations and shows interpopulation differences; while the responses of Uralis and Sholagas to the horizontal–vertical illusion did not differ, those to the Müller-Lyer illusion did: $t\ (26) = 2.4$, $p = 0.02$.

The differences between the responses to the two illusions suggest that they may differ in origin. Rivers speculated that the horizontal–vertical illusion may be physiological and influenced by experience; experience of drawing and geometry would reduce it. In contrast, the Müller-Lyer illusion, he thought, was essentially psychological, affected by the direction of attention.

Thus the results showed that the two illusions were experienced in all the populations tested, but their cogency varied, both with the population and the illusion. The notion that there was an enduring difference between the "savages" and the people of the West appeared therefore not sustainable, nor that "savages" were "all the same."

In spite of their importance, the findings of the Cambridge Expedition did not attract widespread attention. This neglect might have been caused by the appearance of the initial findings in an anthropological publication, but even when they caught the eye of an eminent psychologist, the investigations of illusions were not extensively treated. Woodworth (1910) discussed the report extensively in the course of his presidential address to the Anthropology and Psychology Section of the American Association for Advancement of Science; he praised the studies of colour vision and mentioned the studies of illusions but briefly and rather surprisingly observed that "certain visual illusions in which length or direction of a line is greatly altered by neighborhood of other lines have been tested and found to be present in all races and to about the same degree" (p. 180). Another eminent psychologist, Titchener (1916), in his acidic address to the American Philosophical Society concerned solely with the Torres Straits Expedition, criticized the methods used in investigations of tactile perception and colour vision and ignored entirely the studies of illusions, and by ignoring them he encouraged their oblivion.

No mention of Rivers's seminal studies is to be found in Woodworth's (1938) *Experimental Psychology*, then the dominant text of the discipline, nor are these studies mentioned in Stevens's (1951) or Osgood's (1962) compendia. As a belated apology for these omissions, Kling and Riggs's (1972) edition of Woodworth and Schlosberg's *Experimental Psychology* mentions both the Cambridge Torres Straits Expedition and the Toda studies.

In that barren period, those few experimental psychologists who investigated visual perception cross-culturally devoted themselves to studies of shape constancy with special reference to indigenous art (Beveridge, 1935; Thouless, 1933), the elusive phenomenon of eidetic imagery (Doob, 1966, 1970; Peck & Hodges, 1937), and two phenomena more closely related to perception of the illusion figures: perception of closure (Michael, 1953) and perception of rotating trapezoid (Allport & Pettigrew, 1957). This period was brought to an end by the joint effort of Segall, Campbell, and Herskovits (1963, 1966).

Their study, which is fully reported in Segall et al. (1966), is notable for the clarity with which it defines the aim of the investigation, the manner in which this aim is to be pursued, and the analysis of the data obtained. The hypotheses which the study was intended to evaluate derived from two opposing notions of the vectors shaping perception. These notions, which can be described as those of phenomenal absolutists (all human beings perceive in the same manner) and ofcultural relativists (the manner of perception is determined by cultural experience) can be thought of as those held by experimental psychologists and cultural anthropologists, respectively.

The data were collected by means of a carefully prepared booklet (Herskovits, Campbell, & Segall, 1956/1969), presenting variants of illusion figures, which was administered in accordance with strict instructions either by anthropologists resident in the field or skilled experimenters familiar with the population tested (Segall et al., 1963, 1966).

Sixteen populations were sampled. Most of these were African (six were Bantu speakers; four spoke western African languages, one a Bushman tongue and one a mixed group of mine boys, who probably used the jargon of the mines [*Chilapalapa*]). The non-African populations were represented by two American samples, a sample of Haunnóo Phillipinos, and South African Europeans. Although not equally representative of all the cultures of the earth, the sample (which for most subsamples included both children and adults) is the most extensive hitherto used in cross-cultural investigations of perception.

Three hypotheses based on the ecological validity of cues were postulated. These were

1. *The carpentered world hypothesis.* People living in a carpentered world, a world in which right angles are the commonplace feature of objects, are, as a result of their daily experience, inclined to interpret all angles as projections of right angles. Therefore, people living in a typical Western environment where buildings embody many right angles, just as many other items such as books, bricks, paving stones, and sheets of paper do, are likely, when presented with illusion

figures incorporating angles, to interpret these angles as projections of right angles; in contrast people living in a world of less "carpenteredness"—people whose houses are hemispherical and who encounter few rectangular objects daily—are less likely to do so.

Geometrical similarity of the retinal projections of the real three-dimensional (3D) world and the illusion figure lies at the nexus of the illusory effect. Mutually orthogonal edges of structures project a pattern similar to the Müller-Lyer figure. The fact that such structures cannot, as projective geometry shows, cast patterns corresponding to Mulller-Lyer figures having interfin angles less than a right angle seems, curiously, overlooked, although it does reduce the range of Müller-Lyer figures which can be postulated to be directly affected by carpenteredness.

This geometrical limitation is further augmented by the physical restriction of the observer's experience of the carpentered world. In order to apprehend a projection of a carpentered corner corresponding to a symmetrical Müller-Lyer figure whose fins diverge by 90° (the type of stimulus frequently used). the eye's line of sight has to coincide with one of the lines defining the corner, which is not physically feasible. This reservation applies not only to the line of sight just described but extends to a pencil of lines of sight that it defines and to the corresponding illusion figures. Influence of the carpentered environment on perception of these figures can only come about by extrapolation from experiences associated with lines of sight outwith this pencil. An analogous phenomenon occurs at the other end of the range of geometrically permissible interfin angles.

2. *Foreshortening of receding horizontals hypothesis.* Certain environments, notably those of open plains extending into the distance, present observers with an experience of foreshortening not available to their counterparts living in a dense jungle which curtails their range of sight to a few paces. An illusion figure which can be interpreted as a projection of an object extending into the distance is therefore likely to be experienced more intensely by the former than by the latter.
3. *Symbolizing three dimensions in two hypothesis.* Cultures differ in the extent to which they embrace representations of 3D objects on flat surfaces, either by depiction or by engraving, and the cultures in which such representations are found may differ markedly, as Beveridge (1939) demonstrated, in the manner of representation. Since illusion figures are, *sui generis*, pictures, one would expect them to reflect cultural usage. Since cultural usage has two distinct characteristics—the extent of passive exposure to pictures and the extent to which active creation of pictures in perspective is fostered—the hypothesis is more complicated than the other two, and identification of suitable populations for its verification is particularly difficult as traditional schools of art differ greatly in their styles.

The booklet of stimuli used in the study presented detailed instructions for administration and four figures for testing the subject's comprehension. This was followed by illusion figures reproduced in black and red, and, with the sole exception of the Sander parallelogram and the Poggendorff figures, the lines not quite touching, so that, for example, in the case of the horizontal–vertical illusion, the stem of T did not quite touch the bar.

The illusions explored were (a) Müller-Lyer, (b) Sander parallelogram, (c) horizontal–vertical, (d) perspective drawing, and (e) Poggendorff (see Fig. I.2-2). The stimuli used to evaluate the first three illusions were modified versions of traditional figures. The perspective drawing consisted of two short oblique lines at different angles, so

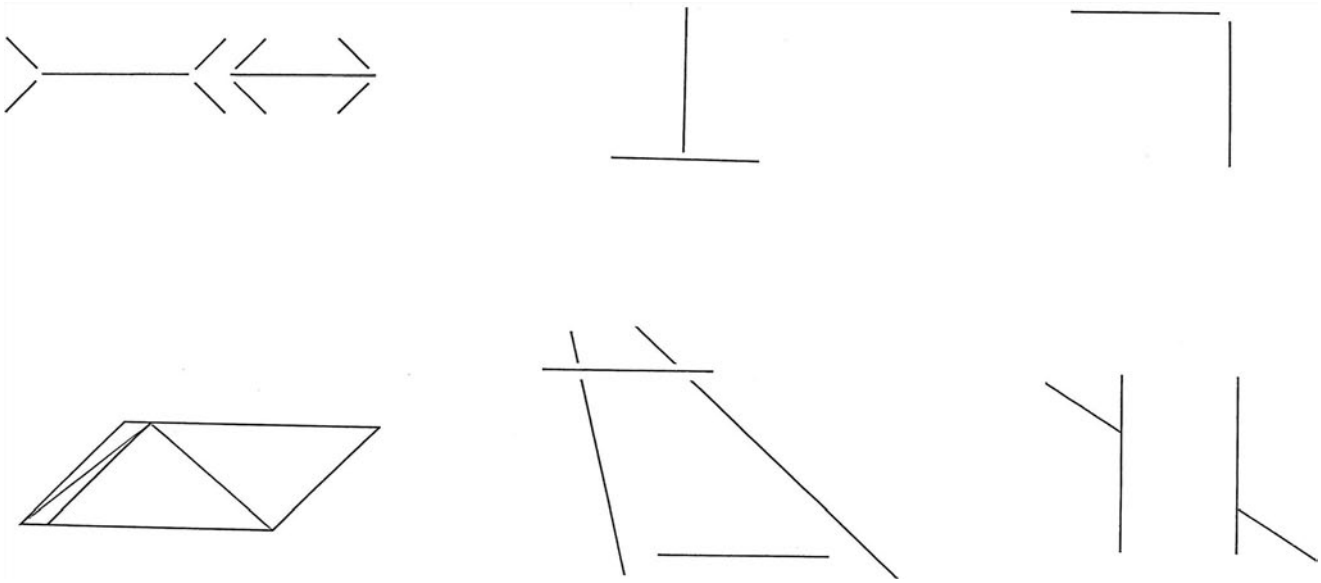


Figure I.2-2. Illusions of the Segall, Campbell, and Herskovits (1969) set. The stimuli were in two colours: black and red. In the case of the horizontal–vertical illusions, the colours of the two elements alternated within the test series. In the case of all other illusions the lines the subject was required to attend to were red. (Adapted from Herskovits et al. 1969.)

that they appeared to converge at some distant point. Two parallel lines, one placed between the converging lines and one above it and cutting across these lines, completed the figure, which can be thought of as a skewed version of the Ponzo illusion. The Poggendorff illusion stimuli consisted of a pair of black vertical lines and red oblique lines. This illusion proved particularly difficult to administer; a very experienced experimental psychologist, thoroughly familiar with the population being tested, failed to obtain reliable data (Reuning & Wortley, 1973). Segall et al. (1966) relegated the responses obtained to the Poggendorff illusion to the appendix of their report. Noteworthy results were obtained with other stimuli.

The samples drawn from more carpentered environments were found to be more prone to the Müller-Lyer and Sander parallelogram illusions than the samples from less carpentered environments with the Müller-Lyer illusion showing stronger effect. Surprisingly, however, the children's responses support the hypothesis more strongly than do those of adults. This seems to be contrary to the hypothesis, since the longer exposure of the adults appears to be associated with the weaker effect. Segall et al. (1966) suggest that this result may be explicable by postulating that subjects learn both to respond unquestioningly to the cue of carpenteredness and to temper the unbridled effects of such responding (this ability is greatly influenced by learning how to represent three dimensions in two). Children have gained the latter skill to a lesser extent than adults. As the results obtained support for the carpentered world hypothesis, they agree with Gregory's (1963, 1966, 2009) notion that inappropriate constancy scaling is responsible for the illusion.

The effect of foreshortening on the horizontal–vertical illusion was assessed using two of the three versions of the figure used by Rivers (1901, 1905): the inverted T and the inverted L. Use of these stimuli ignored Kunnapas's (1955) analysis of the horizontal–vertical illusion, which shows that they may not be equally suitable for investigating the impact of open vistas on perception (see Deręgowski, 1967). Correlational analysis of the responses of all the subjects (N = 1,030) to the two versions yields significant but low product–moment correlation of 0.38, showing that responses of individuals to them are positively correlated. Reuning and Wortley's (1973) comparison of responses to the two illusion figures by 44 Europeans yielded an identical correlation, but an analogous comparison of 100 Bushmen yielded a correlation of zero, suggesting that the results should be interpreted with care. Notwithstanding these observations, the findings presented by Segall et al. (1966) do tentatively support their hypothesis that populations living in open environments experience the illusion more acutely than those living in closed environments.

The Sander parallelogram (see Fig. I.2-3) can be thought of as a compilation of two Müller-Lyer illusions, with the two double-ended arrows having equal shafts but differing in sharpness of their arrowheads: that forming the diagonal of the narrower parallelogram being sharper. Studies of the Müller-Lyer illusion lead one to expect the sharper-ended arrow to be perceived as shorter, and this is generally the case. The parallelogram can also be thought of as an oblique projection of a rectangular surface; such projections are sometimes found in art. The represented rectangle is divided into two unequal rectangles by a line parallel to its edges and normal to the putative observer's fronto-parallel plane. The diagonals of the rectangles are obviously unequal—this too accords with the perceived illusory effect.

Neither of these exegeses is entirely satisfactory since the illusion remains when the stimulus is reduced to two triangles (by deletion of the upper boundary of the parallelogram and of the dividing line; see Fig. I.2-3); they do, however, show that in terms of the hypotheses put forward by Segall et al. (1966), the illusion can be regarded as an embodiment of the Müller-Lyer illusion and hence as a touchstone of the carpentered world hypothesis or as an demonstration of the perspective and therefore as related to the perspective drawing.

The product–moment correlation between the Müller-Lyer and Sander parallelogram scores for all adults (N = 1,030) is reported to be 0.35 and that between the perspective drawing and the Sander parallelogram 0.03. However, as both the Sander parallelogram and the perspective drawing illusions yielded group scores which did not distinguish convincingly among samples, their contribution to cross-cultural studies of perception of illusion is questionable. Segall et al.'s (1966) investigation inevitably provoked further cross-cultural investigations of illusions.

An important advance in cross-cultural studies was made by Gregor and McPherson (1965) using Herskovits's test battery. Its importance lies in the comparison of two groups of similar genetic stock (Australian Aborigines), an aspect entirely ignored by Segall et al. (1966) (An investigation of the same groups by Gimigliano is reported briefly in Segall et al. (1966) as yielding very similar results. However, only Gregor and McPherson's report, which is more detailed, is discussed here.) The two populations, those of Hermannsburg and Yuendumu, inhabited the same "open-visted" environment but differed markedly in their habitats, their ways of life, and the extent of their schooling; the Hermannsburg population was markedly more exposed to "Western" ways including the carpentered world. Accordingly, the Segall et al. (1966) hypotheses would lead one to expect the Hermannsburg population to be markedly more prone to the Müller-Lyer and Sander parallelogram

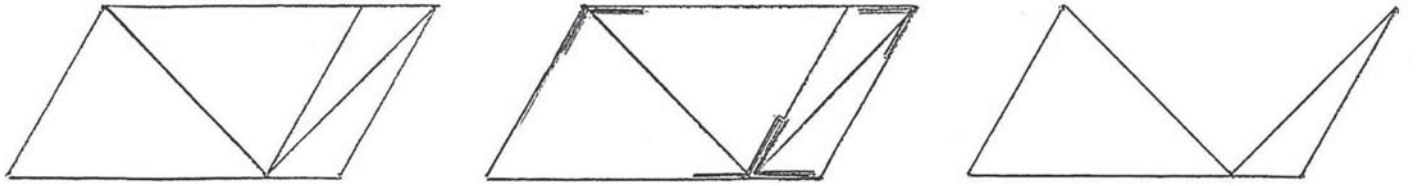


Figure I.2-3. Sander parallelogram, a figure showing its postulated relationship with the Müller-Lyer figure and a figure showing that the illusion persists when the stimulus is reduced to two triangles. (Adapted from Deregowski, 1980.)

illusions and both samples to be about equally, highly responsive to the horizontal–vertical illusion.

The data, which unlike data reported previously were analyzed separately for men and women, do not conform to such a neat scheme. There were gender differences: when responding to the Sander parallelogram the Yuendumu men were more affected by the illusion than the Yuendumu women, but when responding to the horizontal–vertical illusion, Yuendumu women were more affected than Yuendumu men. No analogous difference was observed in the Hermannsburg sample. Thus, while the Yuendumu sample responded in the manner one would expect from Fraisse and Vantrey's (1956) study of responses to the inverted T figure conducted in Paris (adults of low education; art students and science students were used), the Hermannsburg sample did not. The results suggest that the Segall et al. (1966) hypotheses are not infallible in predicting the extent of the illusion and that some other factor, connected with gender, intrudes.

Comparison of the samples derived from the two populations fosters further difficulties for the hypotheses of Segall et al. (1966). Men drawn from the two populations do not differ significantly on either Müller-Lyer or Sander parallelogram illusions, whereas women do so in the predicted direction. This is an unexpected effect, especially since Hermannsburg men's daily work ensures intense exposure to the carpentered environment.

As expected, both versions of the horizontal–vertical illusion affect perception of both groups greatly; there is no significant difference between them, but for Hermannsburg men being more prone to the inverted T stimulus. The proneness to the illusion of the two groups is such that their scores on both versions of the stimulus head the ranked scores obtained by Segall et al. (1966). Among the environments from which samples of the Segall et al. (1966) research were drawn, that of the Kalahari Bushmen corresponds most closely to that of the Australian Aborigines; hence one would expect Bushman scores to lie in close proximity to the Australian scores. This is not so. The Bushman scores fell about at the middle of the scale and were lower than those of the Fang and of the Ijaw; both the Fang and the Ijaw live in dense forests. This discrepancy may be explicable by the surprising perceptual skills of the Bushmen, who in spite of exposure to the environmental press fostering susceptibility to the horizontal–vertical illusion, responded in a manner not differing from that of suburban Evanstonians. Possession of such a skill in turn suggests, as do findings of sex-linked differences by Gregor and McPherson (1965), an innate vector. If so, then the two principal Segall et al. (1966) hypotheses define two putative influences, not determinants of percepts. Reuning and Wortley (1973) succinctly summarize their own studies of a wide range of perceptual processes of Bushmen by saying that, according to their results, Bushmen's illusion susceptibility to the Müller-Lyer and the two versions of the horizontal–vertical illusions falls between those of White and Black South Africans, although Bushmen live in an environment which is most extreme as far as provision of rectangular objects is concerned. It also, as has been said, provides extensive open vistas. The results reported by Morgan (1959) show Bushmen's Müller-Lyer scores as falling between those of Europeans and Black mine laborers, differing significantly from both; Europeans, as expected, were most prone to the illusion. The scores of the three samples on both versions of the horizontal–vertical illusion form the same pattern. There is no significant difference between Bushmen and European scores, which show lower proneness to illusion than that of the Black mine workers.

Mundy-Castle and Nelson's (1962) extensive neuropsychological survey of an isolated community of White forest workers in South Africa offers a comparison between mean illusion susceptibilities of members of that community, urban White adults, and Bushmen. On both Müller-Lyer and horizontal–vertical (both versions) illusions, the forest workers differed significantly from the urban White adults, being less susceptible to the Müller-Lyer illusion and more susceptible to the horizontal–vertical illusion. The susceptibility of the Bushmen was intermediate between those of these two populations, differing significantly from that of the forest workers only in the case of the inverted L form of the horizontal–vertical illusion.

Berry (1968) suggested that some of these confusing results could be clarified by (a) measuring susceptibilities in a single population, thus following Gregor and McPherson's (1965) lead and (b) by taking into consideration perceptual developments of the subjects. The procedure assumes the subject's measure of susceptibility to illusion indicating differences of perceptual development to be independent of experience of carpenteredness. Berry compared responses to Müller-Lyer illusion of samples differing in their experience of carpenteredness, of both Baffin Island Eskimos and Sierra Leone Temne, and reports a modicum of support for his thesis. A question, however, remains whether the measure of perceptual development used by him, Kohs Blocks, was appropriate since the cube can be seen as the very embodiment of carpenteredness, and the Kohs Blocks test involves arranging of cubes.

M'buti Pygmies of the Congo present a population ideally suited to investigation of the effects of absence of carpentered environment on perception of the Müller-Lyer illusion, since not only do they inhabit a forest in which they live by hunting and gathering but their shelters are not merely circular in plan but domed. They are completely illiterate. Samples of M'buti, Bashi (lakeshore agriculturalists living in round huts, who are "largely illiterate"), and Europeans with secondary or higher education (Bonté, 1962) yielded the following results. When a wooden slide similar to that used by Rivers (1901) was used to measure the illusion, all samples showed about the same pronounced proneness. When the illusion was assessed by means of a preliminary version of the Segall et al. (1966) drawings there was a significant difference between the Bashi and the Europeans in the expected sense. The drawings could not be used with M'buti, presumably because of their unfamiliarity with pictures.

The M'buti Pygmies' utter lack of experience of the carpentered world does not seem to have rendered their susceptibility to illusion less than that of either the Bashi or the Europeans. Even more intriguingly, when the Europeans and the Bashi are compared, the effect of the carpentered world is detectable by means of the illusion presented using a drawing but not a slide.

Segall et al. (1966) attempted to deal with the embarrassment of Bonté's (1962) findings by questioning the

procedural aspects of her investigation: (a) the subjects might have misunderstood the instructions (an issue handled with great adroitness by Jahoda [1966] when discussing his Ghanaian findings), or (b) the slide might have led them astray. Three features of the slide are drawn to our attention by the authors as potentially deleterious because (a) the grain of the wood of which the slide is made, (b) the relative shortness of the fins, and (c) the visibility of the edge of the enveloping member of the slide, at which two parts meet. It is this last feature that is thought most likely to have affected the results. It is difficult to evaluate these suggestions, as no hint is offered why these features should increase susceptibility to illusion by the Bashi to the European level.

There appears to be no published data regarding the surface-grain effect—there are several studies on the effect of length of the fins on the illusion, but none appear to address the problem of cross-cultural variation, and the dividing line was present on the slides introduced by Rivers. Indeed, Rivers (1905) used two types of, in one of which the dividing line appears as it did in Bonté's (1962) apparatus (at the tip of the outward pointing arrow) and in another in which it bisected the shaft of the arrow between the inward facing arrows, and compared (Rivers, 1905) data gathered with these instruments from the same population both for the Todas and for the English. They do not differ appreciably. It would also appear somewhat odd if a perceptual norm acquired through persistent exposure were at the mercy of such minor vectors. It is probably safer to treat Bonté's results as yet another question put to the neat hypotheses. Thus it appears that the Segall et al. (1966) hypotheses did kick start cross-cultural studies of illusion again, but they could not alone drive them along.

Jahoda's (1966) work in Ghana confirmed this. He tested those African samples (Lobi and Dagomba, who live in the open country and in a relatively uncarpentered environment, and Ashanti, who live in tropical forests and in a relatively carpentered environment) and a Scottish sample on the Müller-Lyer and the T form horizontal–vertical illusions, both presented in the form of slides. The results obtained were even more surprising than those of Bonté (1962). The two open vistas groups differed radically on the horizontal–vertical illusion, the Lobi mean percentage effect being a mere 1.7, that of Dagomba: 22.7, while the forest dwelling Ashanti scored 16.2. The Scots scored 11.2, differing significantly from both the Lobi and the Ashanti—clearly not a result supportive of the foreshortening hypothesis. The scores obtained with the Müller-Lyer slide also failed to support the relevant carpentered world hypothesis. Although, as one would expect, the grouped African subjects were less prone to illusion than the Scots, there were significant differences among the groups.

Discussing these results, Jahoda (1966) points to the fact that in order for an illusion to occur, the perceiver must recognize that a *picture* represents a shape that has certain spatial qualities. A failure to perceive an illusion may therefore be a consequence of a failure to see the shape in the picture (lack of pictorial skills), or a failure to associate that shape with a particular spatial experience (lack of exposure to such stimuli, open terrain, or carpentered objects, as the case may be), or of course both factors. The former factor will be discussed later.

Confusingly, while Jahoda's (1966) African samples showed no expected effect of carpenteredness, a sample of Navajos tested by Pedersen and Wheeler (1983) did. Even more puzzling results were obtained by Hautaluoma and Loomis (1972) from a sample of boys tested in Nuristan in Afghanistan. They showed greater susceptibility to the Sander parallelogram illusion and lesser susceptibility to the Müller-Lyer illusion than did children from Illinois. The former effect could perhaps be associated with the manner of depiction of rectangular objects (rugs, table tops, pools) in Islamic art of the area. The paper does not address this issue. The Nuristanis live in a carpentered environment. Further studies (Hagen, 1977; Jones, 1974) on samples of populations of European descent living in different environments confirmed that the two hypotheses of Segall et al. (1966) cannot account fully for differences in susceptibility to illusions.

Jones (1974) tested two groups of Canadian schoolchildren drawn from about equally open-visted environments; they differed markedly when compared with children from an urban environment in their scores on the horizontal–vertical illusion, one group supporting the Segall et al. (1966) hypothesis, and the other not at all. As the report does not indicate in what form the illusion was administered, this result cannot be evaluated further. The scores of the same three groups on a test of pictorial depth perception are equally puzzling. One group of rural children (supposedly subjected to fewer cultural stimuli) did indeed score significantly less than the urban children. No such relationship prevails in the case of the other rural group. The group scoring significantly less was the group that did not differ from the urban group on the horizontal–vertical illusion.

Hagen (1977) tested rural and urban American students using the Segall et al. (1966) battery and found (not surprisingly) that the two groups' scores on the Müller-Lyer and the Sander parallelogram did not differ. The results, therefore, tell us nothing of relevance to the carpentered world hypothesis. They do, according to the author, provide partial support for the foreshortening hypothesis, but this assertion is dubious. The support derives from responses to the perspective illusion, the stimulus which failed to discriminate cogently among the Segall et al. (1966) samples (of 15 adult samples, only the Bété yielded truly convincing data). The responses of the urban and rural sample obtained by Hagen to the two versions of the horizontal–vertical illusion differ. Those to the L-shaped figure showed greater susceptibility of the urban sample, and thus contradict the hypothesis, while those to the T-shaped stimulus showed no difference between the samples.

Limitation of cross-cultural investigations to the horizontal–vertical illusion, the Müller-Lyer illusion, and the Sander parallelogram, because these illusions appear to fit neatly to two of the hypotheses put forward by Segall et al. (1966) is deleterious to investigation of illusions in general, because it fails to take into account relationships between illusions, which diverse theoretical speculations suggest. Thus Gregory (2009) groups the Zollner illusion as being the effect of the same vector (rectification of angles) as the Müller-Lyer and thinks that the Poggendorff involves an element of perspective, the oblique line passing behind the barrier of parallel lines. (Both the Poggendorff figure

and the perspective figure failed to yield meaningful results in Segall et al.'s (1966) grand cross-cultural study.) Piaget (1969) groups illusions into primary and secondary, and this grouping, if universally valid, should withstand cross-cultural testing; therefore, for example, the relationship between the Oppel-Kundt illusion and the Müller-Lyer illusion, both primary, should remain the same in all cultures, notwithstanding cross-cultural variation in susceptibility.

Jahoda and Stacey's (1970) study involved such desirable expansion of the range of illusions. The illusions explored were the Helmholtz square, Sander parallelogram, Müller-Lyer, dichosection, horizontal–vertical (L form), Poggendorff, Boring circles, Ebbinghaus (Titchener) circles, Judd 1, Judd 2, Wundt, and Herring. There is no indication how the Segall et al. (1966) hypotheses apply to such a set of stimuli; instead two hitherto unasked questions were asked: Does perception of illusion figures vary cross-culturally between highly schooled samples, and does the nature of schooling affect it? Does training in such disciplines as art and architecture affect it? To answer these questions, samples were drawn from institutions of higher learning in Ghana and Scotland.

In Jahoda and Stacey's (1970) study, the horizontal–vertical was presented in L form only, and a dichosection figure consisting of a tilted inverted T was included in the set to probe the assertion (Deręgowski, 1967) that the horizontal–vertical vector and the dichosection vector should be regarded as distinct. Of the two "weak" figures of the Segall et al. (1966) set, the perspective was not used but the Poggendorff was.

Analyses of the data (parametric and nonparametric) show that the Müller-Lyer (ML) and the dichosection (DS) scores were unaffected by either culture or training. All others (Helmholtz square [HS], Sander parallelogram [SP], horizontal–vertical [HV], Poggendorff [PD], Boring circles [BC], and Ebbinghaus (Titchener) circles [EC]) are affected by both. A correlation matrix of the illusions reveals a spine common to both cultures consisting of those illusions which correlate at the 0.2 level or above and at a significance level of higher than 0.001. These are: DS–ML; ML–SP; SP–HV; and HV–HS. The Ghanaian responses show similar correlations: HS–SP and HS–ML, which are absent from the Scottish responses, whereas the Scottish responses correlate at about the same strength: SP–EC. There are therefore interesting but not readily explicable cultural differences: for Scots, but not for Ghanaians, the EC and SP results correlate; for Ghanaians, but not for Scots, the HS results correlate with both the ML and SP results, thus binding HS to the "spine."

The fact that DS is not affected by either culture or training but HV is affected by both supports the notion that the inverted-T form should not be used to investigate the foreshortening hypothesis. The fact that scores of the *bête noire* of earlier studies, the Poggendorff illusion, do not correlate with any of the other scores confirms its status as does an additional test of untrained Ghanaians, which showed nonsignificant correlations with the Embedded Figures test with all illusions used, with the sole exception of the Poggendorff, which yielded a negative correlation showing that susceptibility decreased with field dependence.

The ecological hypothesis asserts that a percept of a particular visual stimulus is influenced by the circumstances in which perceptually related stimuli were previously encountered; the Müller-Lyer is affected by the experience of the carpentered world. This implies that the setting in which a stimulus is presented may affect its perception. The illusion figures are pictures—and not all cultures are equally familiar with pictures—an additional element may therefore influence cross-cultural differences. Segall et al. (1966) recognized this possibility and suggested that the presence of pictures in a culture may have a dual effect, modulated by the manner of experience of pictures, with passive familiarity increasing and picture making decreasing the illusory effect. This third and rather elusive vector has not, however, been examined empirically by Segall et al. (1966) with a thoroughness comparable to that of carpenteredness and foreshortening. The familiarity with pictures and picture making was rather glibly assessed, a significant lapse since picture making is as culturally widespread as it is variable; even those populations least acculturated to the Western mode of life do have a notion of picture making, as Reuning and Wortley's (1973) studies of Bushman sand drawings shows.

The two principal vectors cannot, evidence shows, adequately account for the observed variations of proneness to illusion. Further vectors therefore need to be put forward. There is an infinite number of candidates, ranging from physiological (such as the pigmentation of fundus oculi) to social (such as the manner of childrearing). The notion that pigmentation may affect perception of illusions, probably because pigmentation is easily assessable, has attracted the most attention.

Pollack (1966) speculated that pigmentation of the fundus may affect perception of illusions, incorporating lines intersecting at angles, and demonstrated this to be so in the case of children's responses to the Müller-Lyer figure (Pollack & Silvar, 1967). This finding together with the report (Silvar & Pollack, 1967) that there are racial differences in pigmentation suggested that, as Pollack points out, any difference of illusion susceptibility between peoples differing in pigmentation might have been affected by this very factor.

This notion was gleefully embraced by Berry (1971), who compared susceptibility to the Müller-Lyer illusion of samples drawn from five different populations. The samples were ranked on their skin pigmentation, this being assumed to offer a reliable measure of their retinal pigmentation. This measure showed distinct ethnic groups to differ surprisingly in the extent of their variation; for example when the two subsamples of Scots and of Eskimos fell into neighboring ranks, the two subsamples from Melanesia differed by five ranks. Analysis of these data led the author to conclude that pigmentation is the significant predictor of Müller-Lyer susceptibility.

A paper by Jahoda (1971) offers no such enthusiastic support. Malawi and Scottish university students responded to blue and red Müller-Lyer figures exposed for about 0.5 s. Responses of the two groups to stimuli of either colour were not found to differ. If it be taken that the Malawi students have experienced carpenteredness to about the same extent as the Scots, then the results show no pigmentation effect. Comparison within the samples show, however, that while

Scots respond to the two colours similarly, Malawians are more affected by red than by blue stimuli. Jahoda (1975) pursued the matter of the effect of colour on perception further using Scots and Ghanaians and found that all subjects did less well with the blue stimuli.

Stewart (1973) explored both the effects of pigmentation and of carpenteredness, using American and Zambian samples. Comparison of White and Black Illinois children whose pigmentation was ophtalmoscopically assessed on Müller-Lyer and Sander parallelogram figures yielded no indication of differences associated with difference in pigmentation.

The Zambian samples were drawn from environments differing in their carpenteredness (and coincidentally in their educational facilities). These samples, which had about equal degree of pigmentation, showed steady increase of susceptibility to illusion with carpenteredness. This finding is contrary to that which Bornstein's (1973) essay on pigmentation and perception of colour and illusions would lead us to expect. The thesis of the essay rests upon Pollack and Silvar's (1967) work already discussed and on a reanalysis of the Segall et al. (1966) data. Although the essay is mostly concerned with the Müller-Lyer illusion, its observations clearly apply to other illusions as well.

Inspired by this essay, Bolton, Michelson, Wilde, and Bolton (1975) carried out an investigation in Peru involving the Müller-Lyer and Sander parallelogram illusions and the two versions of the horizontal–vertical illusion (inverted T and inverted L). They reasoned that since eye pigmentation is affected by ultraviolet radiation, it must be affected by the geographical altitude of one's habitat and the distance from the equator. Two Peruvian populations living at about the same distance from the equator but at markedly different altitudes (1,370 m and 3,960 m) were located. Their manmade environments were about equally carpentered but one inhabited dense forest and another an open plain. This made them ideal subjects for assessment of the effect of pigmentation and experience of foreshortening on perception of illusions.

The findings were that the samples did not differ in their responses to the Müller-Lyer and Sander parallelogram illusions; the pigmentation hypothesis was therefore not supported. They did differ on both versions of the horizontal–vertical, with the inhabitants of the barren plain being more affected, thus supporting the foreshortening hypothesis. It is noteworthy that the difference between mean responses to the two figures obtaining in the two samples are about equal, suggesting that the dichosection effect is about the same and therefore independent of the "open vista" effect. The ocular pigmentation, therefore, although an important factor in perception of illusions, as Coren and Porac (1978) have shown, does not elucidate cross-cultural differences. No other equally attractive candidate to do so has been identified.

Jahoda and Stacey (1970) show no strong relationship between field dependence and illusion perception in the cross-cultural context. An attempt to link perception of illusions to the willingness to abide by instructions, which was made by Davis and Carlson (1970), is clearly unconvincing since no measure of general willingness to abide by instructions on perceptual tasks was used.

As has already been noted, in the course of cross-cultural investigations that grew from Segall et al.'s (1966) seminal study, the number of illusions explored declined. The notable dropouts include the perspective figure, which is of particular interest since the stimulus used is clearly related to the Ponzo figure, one of the best known illusion stimuli. The cross-cultural investigations concerning the Ponzo figure constitute a separate stream, which is considered below.

The convergent lines of the commonly employed Ponzo figure are mutually symmetrical, and the parallel lines they embrace are symmetrically disposed between them. The Segall et al. (1966) perspective figure violates both these rules, and, in addition, the convergent lines are broken, which may reduce their perceptual cogency, and the two sets of lines differ in colour; such colour grouping may affect adversely their mutual influence.

It is possible therefore that the surprisingly nondiscriminatory nature of the perspective figure was due to its idiosyncratic characteristics. This postulate is confirmed by the study of rural, urban, and highly educated Xhosa reported by Smith (1973). (The report contains a misprint of the length of the standard stimulus used.) It was found that the mean extent to which the three groups experienced the illusion varied: the ratio being 51:75:122, and there was a significant difference between the rural and the highly educated Xhosa. Further introduction of a perspective-inducing picture expected to enhance the illusion had a significant expected impact on the educated group but only a trace of such an impact on the urban sample and none on the rural sample.

Analogous results are reported by Leibowitz and Pick (1972), who found no differences between the responses of Ugandan, Pennsylvanian, and Guam students to the Ponzo figure, all of whom showed illusion of about 10% and no illusory effect at all in the case of Ugandan villagers.

A gradual increase of the potency of the pictorial depth cues by use of photographs of an open field and photographs of a railway line, on which the two parallel elements of the Ponzo figure are presented, resulted in the gradual increase of perceived illusion of the Ugandan and Pennsylvanian students; it did not affect the susceptibility of the Guam students. Ugandan villagers consistently showed no susceptibility; their responses to the control stimulus of two parallel lines to the classical stimulus and the two photographs did not differ. The difference between the Guam students and the two other student samples is probably explicable by their distinct visual experience resulting from the environment in which they live (Leibowitz, Brislin, Perlmutter, & Hennessy, 1969). The distinctiveness of the responses of the Ugandan villagers is also likely to result from their visual experience from which pictures are notably absent, a conclusion consistent with the results obtained from the Xhosa samples.

An extensive study of the illusion by Brislin (1974) compared responses of American mainland and Guam subjects to a range of stimuli differing in richness of cues. The two populations differ greatly in their exposure to cues postulated to foster the Ponzo illusion—there are few straight roads and no railways on the hilly island, which therefore differs greatly from Central Pennsylvania whence the American sample came. The ages of the subjects ranged from 3 to 22 years. The stimuli were presented in two

orientations, either with the point of convergence of the two oblique lines being above or to the subject's left. The four variants of the stimulus illusion figure were (a) traditional array, (b) two parallel lines on the background of a depiction of a receding sward, (c) two parallel lines on the background of a depiction of a receding railway track, (d) two parallel lines on the background of a set of 11 lines converging symmetrically to a single point to the observer's left (the problem of orientation of the stimulus was therefore introduced) and a control figure consisting of two parallel lines. The cultural difference was found to be present in the cases of the two stimuli wherein the parallel bars were presented on depictions of receding ground. Surprisingly, a significant difference was also found in the case of the control figure composed of two parallel bars but was absent in the case of the classical figure.

In further pursuit of the problem (Brislin & Keating, 1976), the illusion was presented to three groups of subjects coming from the mainland of the United States, from islands of the Pacific, and from Philippines. They responded to the arrangements of boards laid in front of them on the ground and so arranged as to form a Ponzo figure with its vertex pointing to the left. This was an important innovation; testing of an illusion effectiveness in real space should, Brislin and Keating suggest, strengthen the effect. The results were consistent with the foreshortening hypothesis; both the mainland and the Philippine sample showed susceptibility of about 10%, about double that of the islanders. The result is puzzling, however, since susceptibilities of US adults responding to a picture of railway tracks with apex to the left, determined earlier (Brislin, 1974), are more than twice those determined in real space. A depiction of space is more spatial than the space itself.

The traditional Ponzo figure consisting of two converging lines and two bars presents a single depth-related cue (convergence) on a surface of paper showing no depth. When the surface of the paper is embellished with a texture gradient, an additional depth cue is introduced which may be consonant with that provided by the convergent lines or otherwise. Consonance, Leibowitz et al. (1969) show, strengthens the illusion. (There appears to be no empirical evidence of the combined effect of the convergent lines, with the texture gradient depth cues being mutually orthogonal.) Such orthogonality occurs when, as in the case of the Brislin and Keating (1976) study, a leftward-pointing illusion figure is presented on the ground. In addition to these two vectors there is also a vector resulting from the very three-dimensionality of the arrangement and therefore from binocular perception of true depth. All the samples referred to above show susceptibility to illusion. A significant deviation from that rule is provided by the sample of Ugandan villagers, who showed no susceptibility to either the geometric figure or the figures incorporating texture gradients (sward and rail track; Leibowitz & Pick, 1972).

It is possible that the Ugandan villagers of the Leibowitz and Pick (1972) study were consistently immune to the illusion because of their unfamiliarity with pictures, which ensured their percepts were dominated by the flatness of the surfaces on which the stimuli were presented, and ignored the pictorial depth cues the depictions contained. This problem was examined by Kilbride and Leibowitz (1975). They segregated Ugandan villagers on the basis of their responses to two of the pictures used by Leibowitz and Pick: (a) those who recognized correctly or approximately correctly both depictions, (b) those who did so in the case of one picture only, and (c) those who called them by entirely wrong names (e.g., calling a railway; a ladder or an open field a roof). Responses of the three groups of subjects to the Ponzo illusion were markedly different; those in the first group were similar to those of the Ugandan students in Leibowitz and Pick's study. Thus, it is argued, subjects whose percepts were more affected by the physical surface (as shown by the fact that their responses to the introductory pictures implied small pictorial depth) were less prone to the illusion. One would expect such subjects to be also less prone to the horizontal–vertical illusion and their susceptibilities to the two illusions to correlate positively. There appears no extensive relevant published evidence. Product–moment correlation obtained by Segall et al. (1966) between the inverted-L horizontal–vertical and their perspective figure is a mere 0.03, and not the significant positive value one would expect.

The Ebbinghaus (Titchener) circles illusion which Jahoda and Stacey (1970) found to be perceived differently by Scottish and Ghanaian subjects (Scots being more prone, and training increasing proneness in both samples) has recently attracted the attention of a group of researchers working with the Himba, a remote cattle-keeping people of northern Namibia. The Himba's susceptibility was found to be less than that of the English (de Fockert, Davidoff, Fagot, Parron, & Goldstein, 2007); thus the result was analogous to that reported by Jahoda and Stacey. Further, when modified stimuli were used in which the circular-inducing elements were replaced by squares, the illusion was reduced but the relative lesser susceptibility (greater accuracy) of the Himba remained. In an extension to this investigation (Davidoff, Fonteneau, & Goldstein, 2008), Ebbinghaus figures composed of silhouettes of cows were used. This reduced the illusion perceived by both groups, but the reduction was markedly greater in the case of the English subjects. The lesser susceptibility of the Himba is thus shown to be a robust effect, and the effect, as Jahoda and Stacey have noted, is puzzling. Since the illusion and the Ponzo illusion share the attribute of detached elements affecting perception of central elements (in the case of the former it is the effect of a ring of inducing elements upon the core element, and in the case of the latter it is the effect of oblique lines upon the parallel segments), one would expect a positive correlation between these illusions. Unfortunately, Jahoda and Stacey did not include the Ponzo illusion in the set of stimuli used in Ghana, nor, it appears, was it used in Namibia.

The effectiveness of several illusion figures, most notably those of Müller-Lyer, Sander parallelogram, horizontal–vertical, and Ponzo, is thought to be associated with illusory perception of depth in a pattern presented on a flat surface, that is to say, with pictorial perception; therefore, cultural differences in responses to pictures need to be briefly addressed. This need is sustained by the difference between responses to the Oppel-Kundt figure obtained by Vurpillot (1963) whose stimuli incorporated hints of the third dimension but whose illusion was presented

on a plane (mean susceptibility observed 2.8%) and those obtained by Deręgowski and McGeorge (2006) to truly 3D stimuli (mean susceptibility observed 5.5%).

Recognition that a pattern of a pigment on a surface is a depiction may involve either one or both of the following attributes:

1. *Epitomic*, which leads to perception of a similarity between the pattern and some object but does not involve perception of depth in the picture. This occurs, for example, when a monochromatic silhouette of a head in profile is recognized as representation of a head. The epitomic effect is of little relevance to perception of illusions and is not examined at length here. It should, however, be noted that there appears to be no consensus as to its influence in remote populations. Herskovits (1959) describes graphically a Bush Negro woman's failure to recognize a photograph of her own son, but Rivers (1905) reports that for the Todas "looking at portraits of themselves" was "an occupation of which they never tired." Deręgowski, Muldrow, and Muldrow (1972) report difficulties consistent with Serpell and Deręgowski's (1980) analysis of the process of pictorial perception as a skill. They describe a slow but purposeful search for meaning of clear pictures printed on cloth by the Me'en of Ethiopia, whose culture is pictureless.
2. *Eidolic*, which evokes perception of depth, although the pattern may not resemble any known object. This occurs, for example, when such impossible figures as the two-pronged trident is viewed and is also responsible for the "reversals" of the Necker cube.

It is noteworthy that such reversals are affected by yet another type of depth cue, which is seldom considered and which could be called ethological—the eyespots. Dziurawiec and Deręgowski (2007) have shown that placement of eyespots on depicted animals affects perceived direction of their heading, so that the animals are judged to pass closer to the observer. Consistent with this finding, it has been shown that eyespots placed on one of the elements of a reversible figure, that alternate in their apparent proximity to the observer, cause that element to be perceived as being closer to the observer for a larger proportion of the time of inspection (Deręgowski & Gray, 2013).

It is possible to see in an epitomic figure a portrayal of a 3D object, although the figure lacks any perceptual hint of three-dimensionality, and it is possible to see a third dimension in an eidolic figure even when the figure is self-contradictory (see Fig. I.2-4). Most pictures embody both attributes, but it is the second effect which is of interest in the case of illusions.

Cross-cultural differences in the perception of pictorial depth (the eidolic effect) were cogently demonstrated by Hudson's (1960) study. As the study and extensive work it inspired are described by Miller (1973) and Deręgowski (1980, 1989, 1991), only general observations are presented here.

Although pictorial depth is easily accessible conceptually, the measurement of its perception presents a variety of problems, because both the stimulus presented and the nature of the requested response determine the subject's

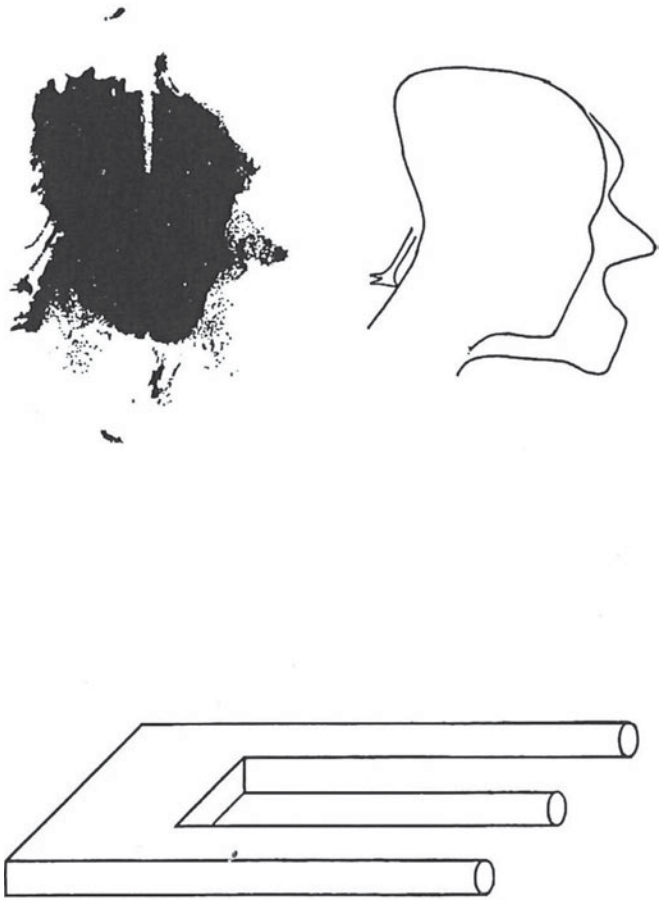


Figure I.2-4. Top: An unintentionally epitomic inkblot and outlines of three of the plethora of its possible percepts. Below: Eidolic two-pronged trident (after Deręgowski, 1984).

performance. The pictures used as stimuli may present any combination of the eidolic and epitomic cues, and the responses made may vary from oral descriptions of depictions to 3D reproductions of the depictions. The differences such variations may bring about are illustrated by the following: (a) Omari and McGinitie (1974) used Hudson's original figures and their modifications with appropriately modified questions to which subjects responded orally and found a significant difference between the two procedures; (b) Deręgowski (1968) administered Hudson's test and a task which required construction (using clay and bamboo splints) of depicted simple geometric models and found that subjects who showed no evidence of perception of the pictorial space on Hudson's test did so on the construction task; (c) Jahoda and McGurk (1974), who evaluated several measures of perception of pictorial depth, proposed a new performance measure in which puppets representing a woman and a girl were used. The test scores on the new test were higher than those on Hudson's test both for Scottish and Ghanaian children, but the scores of the Ghanaians on Hudson's test were at chance level.

It is also noteworthy that identity of environment does not always ensure similarity in pictorial depth perception. Thus Deręgowski and Bentley (1986) requested two groups of Bushman children from a remote area, the!Xe (Sekele) and the Kxoe (Kwengo), as well as a group of urban Zulu

children, to set callipers at an angle of a depicted set and found that the two groups of Bushmen differed significantly and that none of them differed from Zulus, who showed intermediate pictorial depth perception.

Such callipers offer a subtler manner of assessing the impact of a carpentered world experience than do illusion figures. One would expect subjects from a carpentered environment to (a) to set the callipers at a right angle in response to pictures of callipers set at a right angle regardless of the true angle of the depicted callipers *in the plane* of the picture; (b) to set callipers at an angle closer to the right angle than the *depicted* angle in response to a picture showing callipers set at an obtuse angle; and (c) when responding to pictures of callipers set at an acute angle, to comply with the restrictions considered in discussion of the carpentered world hypotheses and therefore to show no drift toward right-angularity. These expectations are supported by the responses of a small sample of students from a Scottish university (Deręgowski, 2013).

The intracultural differences in pictorial depth perception are therefore likely to constitute an important, if ill-explained, influence on cross-cultural measures of susceptibility to illusions. There appears to be no readily accessible data pertaining to the relationship of pictorial depth perception to illusion susceptibility in the most often investigated, Western cultures.

AGE AND GENDER

The impacts of the two socially important individual attributes, age and gender, on perception of illusions in various cultures have been investigated only sporadically. Thorough investigation of these attributes may indicate how distinct manners of socialization and distinct social roles affect perception.

Segall et al. (1966) report that generally the children they tested were more prone to illusions than adults. This is indeed so, but further examination of their data pertaining to the four most salient illusions—Müller-Lyer, Sander parallelogram, and two types of horizontal–vertical—shows that in no culture were children more susceptible to all four of them, and the greater susceptibility of children was significantly more pronounced in the case of the first pair (postulated to be affected by carpenteredness) than of the second pair (postulated to be related to perception of foreshortening). Indeed, in the latter case only in one culture, the Hanunóo of Philippines, were children consistently more susceptible than adults.

These data do not support a simple theory that susceptibility is directly proportional to experience, nor that the age-related changes in susceptibility progress in the same fashion for both pairs of illusions. Nor do they suggest what factors determine susceptibility in any particular culture. The data suggest that in the case of the Müller-Lyer and Sander parallelogram illusions, the developmental process, whether affected by experience or by some physiological changes, must be such that an age of maximal susceptibility occurs before adulthood. In the case of the horizontal–vertical illusion, the data do not offer such a suggestion as shown by the ambivalence of responses of 10 out of 11 samples of children. For the perspective figure, higher susceptibility of children is reported in seven, of adults in three, and of neither in one case.

Elucidation of these observations is only attainable by means of developmental curves describing changes of susceptibility in various cultures over a span of years. Such comprehensive data are not available, but there are scattered items of data relevant to the issue.

Wagner (1977) measured susceptibility of the Ponzo effect using the stimuli previously used by Leibowitz et al. (1969): an abstract figure, an open sward, a railway line, and a control figure. Four age cohorts were tested (7, 10, 14, and 19 years, all boys) drawn from urban and rural Morocco from the schooled and the unschooled. The results show consistent decline of susceptibility with age for all groups when the abstract figure was used. For the two figures presenting texture gradients, the results are less consistent. There is a clear tendency for the susceptibility of the schooled groups to increase with age and for the unschooled rural sample to remain roughly constant through the age range. Susceptibility of the unschooled urban population to the two stimuli differs; it appears to increase with age in the case of the stimulus incorporating rail tracks and not so in the case of the stimulus showing only sward. The rural unschooled subjects present no such clear trends. Wagner argues that these results can best be understood in terms of Piaget's (1969) notions of primary and secondary illusions.

Brislin (1974), using the same stimuli, measured responses to the Ponzo figure of subjects aged between 3 and 18 drawn from the mainland United States and from Guam. Comparison of his findings with those of Wagner within the age range explored by Wagner shows a similar decline for the abstract figure and a clear disparity of trends between the two cultural groups; the susceptibility of the mainland sample increased and that of the Guam sample remained stable at about 12% in the case of the sward and 17% in the case of the rail tracks. Thus the responses of the Guam subjects show the same pattern as those of the unschooled Moroccans. Comparison between responses to the sward stimulus with those to the rail track stimulus provides a measure of the impact of the Ponzo illusion; the data of both studies suggest that rail track stimuli increase the susceptibility scores throughout the age range.

Ahluwalia (1978) measured susceptibility to the Müller-Lyer figure (and to a similar figure in which arrowheads were replaced by rings) of Zambian subjects aged 8 to 14 and 17 to 19 drawn from carpentered and uncarpentered environments. The interpolated curve between responses to the Müller-Lyer figure obtained in both environments fell from 13% to 7% of illusion. Stewart's (1974) research contrasted Zambian and Illinois subjects aged between 6 and 17. Her innovative measure was a miniature Ames Room, a 3D device projecting an image whose perceived shape depends on the extent of the assumption of the rectangularity (carpenteredness) of the visual stimulus (Gregory, 2009; Kilpatrick, 1961). All Stewart's samples show the same age trend; the susceptibility to illusion decreases with age. The proneness to illusion of the Zambian subjects (who included unschooled Valley Tonga) differs at each age level; Zambians are less affected by the illusion than are the Illinois Whites. (A device related to the device used by Stewart, a rotating trapezoid, was used by Allport

and Pettigrew [1957] in a cross-cultural investigation comparing urban Africans, rural Africans, and Europeans. Since it involved perception of motion, it falls outside the scope of this review.)

An investigation involving changes of susceptibility to the horizontal–vertical illusion on a population whose daily tasks involved estimation of distances and therefore of foreshortening was carried out on Costa Rican fishermen by Pollanac (1977). The subjects' experience of sea fishing, and hence of judgement of distances over open expanses of the sea, was thought likely to affect their responses to the horizontal–vertical illusion. Two versions of the illusion stimuli were used: the inverted T and the backward L. Correlations of illusion scores with (a) years of fishing, (b) status on the boat, and (c) age were found to be significant and positive in the case of the inverted-T stimulus, but no such correlations were found in the case of the backward L. The support these results offer to the foreshortening hypothesis is questionable because it has been shown (Jahoda & Stacey, 1970) that the two figures do not merely present different versions of the same illusion but different illusions, and here the inverted-L figure presents the postulated effect in a purer form, yet it does not correlate with the extent of the relevant experience.

The most extensive cross-cultural investigation of developmental changes in susceptibility to four of the illusions—Müller-Lyer, Sander parallelogram, and the two forms of the horizontal–vertical—is that of Dawson, Young, and Choi (1973). The subjects were Hong Kong Chinese, aged from 3 to 21. Interpolated curves of susceptibility against age have all essentially the same shape. The high initial susceptibility at the age of 3 declines at a decreasing rate until a horizontal asymptote is reached. The shape of each of the four curves can therefore be fairly defined by the score at the age of 3, the age at which the asymptote is reached and the susceptibility score is defined by the asymptote. These measures given in this order are

Müller-Lyer: 7.5; 16 years; 4.3.
Sander parallelogram: 5.2; 10 years; 3.5.
Inverted-T horizontal–vertical: 7.2; 10 years; 4.8.
Inverted-L horizontal–vertical: 7.5; 10 years; 5.6.

The most striking characteristic of these data is the difference between the asymptotic susceptibilities of the two horizontal–vertical illusions. The reports from other cultures consistently show greater susceptibility to the T form than to the L form, but here the reverse holds. As the susceptibility at 3 years to both illusions is about the same and the asymptote is reached at about the same age, the perceptual change, which for whatever reason takes place between the ages of 3 and 10, is more marked in the case of the T form. There is no immediately apparent reason why this should be so.

Cross-cultural investigations of age trends are confined to very few illusions, but the illusions investigated are those most often studied. The most striking result of the investigation is the contrariness of the trend between the sensibility to the Ponzo illusion (which increases with age) and sensibility to the four illusions investigated by Dawson et al. (1973) (which declines with age). Since it is often speculated (e.g. Gregory, 2009) that both the Ponzo and the horizontal–vertical illusions are greatly affected by experience (and this is supported by intergroup differences in Wagner's [1977] study), this contrariness calls for an explanation of why experience affects them differently.

Since the social roles fulfilled by men and women vary among societies, and since such roles impose distinct activities which have to be learned and thus various perceptual skills acquired, it is possible that in addition to the biological vector influencing perceptions of illusions by the two genders there may also exist a "social" vector which causes the pattern of differences in perception between men and women to vary with culture. Although this argument can only be verified empirically by cross-cultural studies, it has not gained widespread attention. Studies linking social structure to susceptibility to illusions have not been performed, probably because categorization of societies according to their social structure (bearing in mind those aspects of social structure which are likely to affect illusion susceptibility) is inherently very difficult. It is much more difficult than categorization of environments as carpentered or otherwise, or as offering extended vistas or otherwise.

Only one study, that of Gregor and McPherson (1965), discussed above, addresses the issue at some length. It is important to note that their results show significant differences between susceptibilities of men and women, which vary both with the population and with the illusion stimulus. Unfortunately, Dawson et al.'s (1973) developmental study of Hong Kong Chinese fails to clarify the relationship between illusion susceptibility and gender. Thus he reports greater susceptibility of girls and women as occurring (a) for Müller-Lyer at the ages of 5 and 21 and (b) for inverted L horizontal–vertical at the ages of 3, 12 to 15, and 21 years. No significant intergender differences are reported at other ages within the 3 to 21 age range, and none are reported within that range for Sander parallelogram and inverted-T stimuli.

The complexity of the problem is also made apparent by the results obtained by the same experimenter from the same subjects using different methods. Stewart (1973) found females to be less susceptible than males to both Müller-Lyer and Sander parallelogram illusions; when the same subjects were tested using a miniature Ames Room, the relationship was reversed, with females being more susceptible; the latter result is compatible with that obtained by Deręgowski, Shepherd, and Slaven (1997) on a task requiring judgement of distances in real space. The incompatibility of responses obtained by Stewart using illusion figures with those obtained using 3D models appears to be of the same nature as that prevailing between Vurpollot's (1963) and Deręgowski and McGeorge's (2006) observations of the Oppel-Kundt illusion, both indicating that pursuit of illusions in 3D space presents special difficulties. Further, Stewart reports that while females were more susceptible in five of six cultural samples, they were not in the case of schooled rural Tonga. It should be noted that Jahoda's (1979) study comparing Scottish and Ghanaian primary schoolchildren on Piagetian tasks found the same pattern of sex differences in both cultures. He also found responses to depictions of stimuli to be weaker in the Ghanaian children; their responses to solid models did not differ from those of the Scots.

CONCLUSIONS

The historical progress of cross-cultural studies of optical illusions is uneven. Interest in such comparisons varies with time, as do the motives for engaging in them. The fundamental questions Rivers and his colleagues addressed were: "Are *they* different from *us*? Do *they* perceive the world differently?" Illusion figures were used as tools to answer these questions—tools no different in principle from Holmgren's wools used to assess colour vision or Snellen's chart for assessing visual acuity.

The second wave initiated by the work of Segall et al., which followed more than half a century later, posed an entirely different and more precise question, namely: "Do the differences of visual experience brought about by the environment (natural and cultural) in which individuals live affect their manner of seeing? Does the environmental pressure containing elements similar to an illusion figure affect the percept of that illusion?"

Thus from mere tools that are themselves of little interest, the illusions have gained significance. This significance is further enhanced by their importance to neuropsychological studies, particularly to the effects of laterization (Medvedev & Shoshina, 2004; Rasmjou, Hausmann, & Güntürkün, 1999).

Two principal hypotheses were discussed, both resting on the assumption that experience affects susceptibility of illusions. The carpentered world hypothesis postulated that susceptibility to the Sander parallelogram and the Müller-Lyer illusions would increase with experience of rectangular objects. The "foreshortening of receding horizontals" hypothesis postulated that susceptibility to the horizontal–vertical illusion would be greater in those observers who in the course of their daily experience had learned to associate vertical lines projected on their retinas with objects receding away from them rather than with upright objects. There was also a hypothesis concerning the moderating role of depictions and nondepictive patterns found in a cultural milieu, which was not, however, unlike the other two, pursued with vigour.

The hypotheses were tested on the largest number of subjects and the most varied sample ever used in cross-cultural investigations of perception. Responses to the Müller-Lyer, Sander parallelogram, and horizontal–vertical stimuli supported, with some reservations, the first two hypotheses, but responses to the perspective figure and the Poggendorff were uninterpretable. The study spawned several partial replications on previously untested populations and also provided an inspiration for studies in which illusions other than those linked directly to the hypotheses were used. Such was Jahoda and Stacey's (1970) study, which compared Scots and Ghanaians on seven additional illusions. Scots were generally found to be more susceptible to illusions than the Ghanaians. Results pertaining to two of the illusions need to be noted. The results obtained with the Ebbinghaus circles anticipate and are congruent with those obtained by Davidoff et al. (2008) from the Himba, and absence of correlations between responses to the Poggendorff illusion and those to any of the others confirms its singular character.

Further studies of the Himba responses to the Ebbinghaus circles (Caparos et al., 2012) show that rural Himba are less prone to this illusion than urban Himba, urban British, and urban Japanese observers and that seemingly minimal contact with an urban environment markedly increases susceptibility to the illusion of rural Himba. This finding is difficult to reconcile with Jahoda and Stacey's (1970) observation that two full years or more of systematic instruction in drawing or related skills did not affect responses of Ghanaian students to the illusion, their susceptibility remaining lower than that of their Scottish counterparts.

Cross-cultural differences in susceptibility to such an array of illusions cannot be readily understood in terms of the two hypotheses advanced by Segall et al.(1966); other cultural or perhaps environmental effects have to be identified, or, alternatively, an element of perceptual process not involving changes in the perceptual organization of stimuli but moderating the form in which they reach the cortex needs to be identified. Pollack (1966, p. 82) proposed that the true developmental processes which "reflect quantitative changes in the underlying organising process that determine perceptual behaviour" are distinct from the pseudo-developmental which involves quantitative "changes of peripheral receptor system," which include increase of retinal pigmentation. This ensures that individuals with greater density of fundus oculi are more prone to illusions involving convergent lines than those with lesser density. Coren and Porac (1978) using Caucasian subjects showed it to be so in the case of the Müller-Lyer illusion and not so in the case of the Ebbinghaus illusion, which does not involve such lines. This, if applicable in cross-cultural settings, would clarify some of the findings. However, Jahoda and Stacey's (1970) comparison between Ghanaian and Scottish subjects question such a notion; Ghanaians and Scots do not differ in their responses to the Müller-Lyer illusion.

The ways of cultural influence are sometimes obscure. Ishii, Tsukasaki, and Kitayama (2009) have recently conducted an experiment in which subjects were required to recognize depicted objects of which only a small segment or segments was shown or to recognize it when the entire object was shown but obfuscated. There was a difference between the European Americans and Japanese as well as the Asian Americans. Although the three groups did not differ in their responses to the holistic stimulus, the responses of the European Americans to the partial stimuli were significantly more accurate. The authors speculate that this cross-cultural difference may be a consequence of socialization. It is not clear whether this socialization is expected to affect the process of perception—in which case it would be of interest to students of illusions—or whether it is expected to affect such factors as the willingness to take a risk or fear of making a wrong response—in which case it is of little interest to them.

The great wave of research activity concerning cross-cultural differences in perception of illusions has now, as the list of references shows, abated. This does not mean that all the questions have been answered. It may mean that the value of such studies to understanding of perceptual processes is underestimated.

ACKNOWLEDGMENTS

The author is indebted to Prof. Gustav Jahoda for his comments on the draft of this chapter.

REFERENCES

Ahluwalia, A. (1978). An intra-cultural investigation of susceptibility to "perspective" and "non-perspective" spatial illusions. *British Journal of Psychology*, *69*(2), 233–241.

Allport, G. W., & Pettigrew, T. F. (1957). Cultural influence on the perception of movement: The trapezoidal illusion among Zulus. *Journal of Abnormal and Social Psychology*, *55*, 104–113.

Berry, J. W. (1968). Ecology, perceptual development and the Müller-Lyer illusion. *British Journal of Psychology*, *59*(3), 205–210.

Berry, J. W. (1971). Müller-Lyer susceptibility, culture ecology or race? *International Journal of Psychology*, *6*(3), 193–197.

Beveridge, W. M. (1935). Racial difference in phenomenal regression. *British Journal of Psychology*, *26*, 59–62.

Beveridge, W. M. (1939). Some racial differences in perception. *British Journal of Psychology*, *26*, 59–62.

Bolton, R., Michelson, C., Wilde, J., & Bolton, C. (1975). The heights of illusion: On relationship between attitude and perception. *Ethos*, *3*(3), 403–424.

Bonté, M. C. (1962). The reaction of two African societies to the Müller-Lyer illusion. *Journal of Social Psychology*, *58*(2), 265–268.

Bornstein, M. H. (1973). The psychophysiological component of cultural difference in color naming and illusion susceptibility. *Behavior Science Notes*, *8*(1), 41–101.

Brislin, R. W. (1974). The Ponzo illusion: Additional cues, age orientation and culture. *Journal of Cross-Cultural Psychology*, *5*(2), 139–161.

Brislin, R. W., & Keating, C. F. (1976). Cultural differences in the perception of a three-dimensional Ponzo illusion. *Journal of Cross-Cultural Psychology*, *7*(4), 397–412.

Caparos, S., Ahmed, L., Bremner, J., de Fockert, J., Linnell, K. J., & Davidoff, J. (2012). Exposure to an urban environment alters the local bias of a remote culture. *Cognition*, *122*(1), 80–85.

Coren, S., & Porac, C. (1978). Iris pigmentation and visual-geometric illusions. *Perception*, *7*, 473–477.

Davidoff, J., Fonteneau, E., & Goldstein, J. (2008). Cultural differences in perception: Observations from a remote culture. *Journal of Cognition and Culture*, *8*, 189–289.

Davis, C. M., & Carlson, J. A. (1970). A cross-cultural study of the strength of the Müller-Lyer illusion as a function of attentional factors. *Journal of Personality and Social Psychology*, *16*(3), 403–410.

Dawson, J. L. M., Young, B. M., & Choi, P. P. C. (1973). Developmental influences on geometric illusion susceptibility among Hong Kong Chinese children. *Journal of Cross-Cultural Psychology*, *4*(1), 49–74.

de Fockert, J., Davidoff, J., Fagot, C., Parron, J., & Goldstein, J. (2007). More accurate size contrast judgements in the Ebbinghaus illusion by a remote culture. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 738–742.

Deregowski, J. B. (1967). The horizontal–vertical illusion and the ecological hypothesis. *International Journal of Psychology*, *2*(4), 269–273.

Deregowski, J. B. (1968). Difficulties in pictorial depth perception in Africa. *British Journal of Psychology*, *59*(3), 195–204.

Deregowski, J. B. (1980). *Illusions, patterns and pictures.* London: Academic Press.

Deregowski, J. B. (1984). *Distortion in art: the eye and the mind.* London: Routledge and Kegan Paul.

Deregowski, J. B. (1989). Real space and represented space: Cross-cultural perspectives. *Behavioral and Brain Sciences*, *12*, 51–74, 98–119.

Deregowski, J. B. (1991). Espacio real y espacio representado: Perspectives transculturales. *Estudios Psychologicos*, *1*, 23–118.

Deregowski, J. B. (2013). On the Müller-Lyer illusion in the carpentered world. *Perception*, *42*, 790–792.

Deregowski, J. B., & Bentley, A. M. (1986). Perception of pictorial space by Bushmen. *International Journal of Psychology*, *21*(4), 743–752.

Deregowski, J. B., & Gray, C. (2013). Eyespots in the reversible setting. *Perception*, *42*, 248–250.

Deregowski, J. B., & McGeorge, P. (2006). Oppel-Kundt illusion in three-dimensional space. *Perception*, *35*, 1307–1314.

Deregowski, J. B., Muldrow, E. S., & Muldrow, W. F. (1972). Pictorial recognition in a remote Ethiopian population. *Perception*, *1*, 417–425.

Deregowski, J. B., Shepherd, J., & Slaven, G. A. (1997). Sex differences on Bartel's task. *British Journal of Psychology*, *88*, 637–651.

Doob, L. W. (1966). Eidetic imagery: Cross-cultural will-o'-the-wisp? *Journal of Psychology*, *63*, 13–34.

Doob, L. W. (1970). Correlates of eidetic imagery. *Journal of Psychology*, *76*, 223–230.

Dziurawiec, S., & Deregowski, J. B. (2007). The perceptual effect of eyespots; or the devil may seem closer than she really is. *Australian Journal of Psychology*, *59*, 101–107.

Fraisse, P., & Vantrey, P. (1956). The influence of age, sex and specialized training on the vertical–horizontal illusion. *Quarterly Journal of Experimental Psychology*, *8*, 114-120.

Gregor, A. J., & McPherson, D. A. (1965). A study of susceptibility to geometric illusion among cultural subgroups of Australian Aborigines. *Psychologia Africana*, *11*, 1-13.

Gregory, R. L. (1963). Distortion of visual space and inappropriate constancy scaling. *Nature*, *199*, 678–690.

Gregory, R. L. (1966). *Eye and brain.* London: Weidenfeld & Nicholson.

Gregory, R. L. (2009). *Seeing through illusions.* Oxford: Oxford University Press.

Haddon, A. C. (1901). Preface. In A. C. Haddon (Ed.), *Reports of the Cambridge Anthropological Expedition to Torres Straits* (p. 1). Cambridge, UK: Cambridge University Press.

Hagen, D. J. (1977). Intra-cultural differences in the perception of visual illusions: An urban/rural comparison. *Perceptual and Motor Skills*, *45*, 1105–1106.

Hautaluoma, J. E., & Loomis, R. J. (1972). Perception of visual illusions in a sample of Afghan boys. *Journal of Social Psychology*, *87*(1), 143–144.

Herskovits, M. J. (1959). Art and value. In R. Redfield, M. J. Herskovits & G. F. Ekholm (Eds.), *Aspects of primitive art* (pp. 42–68). New York: Museum of Primitive Art.

Herskovits, M. J., Campbell, D. T., & Segall, M. H. (1956). *Materials for cross-cultural study of perception.* Indianapolis: Bobbs-Merrill.

Herskovits, M. J., Campbell, D. T., & Segall, M. H. (1969). *A cross-cultural study of perception.* Indianapolis: Bobbs-Merrill.

Hicks, G. D., & Rivers, W. H. R. (1908). The illusion of compared horizontal and vertical lines. *British Journal of Psychology*, *2*(3), 243–260.

Hudson, W. (1960). Pictorial depth perception in sub-cultural groups in Africa. *Journal of Social Psychology*, *52*(2), 183–208.

Ishii, K., Tsukasaki, T., & Kitayama, S. (2009). Culture and visual perception: Does perceptual inference depend on culture? *Japanese Psychological Research*, *51*, 103–109.

Jahoda, G. (1966). Geometric illusions and environment: A study in Ghana. *British Journal of Psychology*, *57*(1–2), 193–199.

Jahoda, G. (1971). Retinal pigmentation, illusion susceptibility and space perception. *International Journal of Psychology*, *6*(3), 199–208.

Jahoda, G. (1975). Retinal pigmentation and space perception: A failure to replicate. *Journal of Social Psychology*, *97*(1), 133–134.

Jahoda. G. (1979). On the nature of difficulties in spatial-perceptual tasks: Ethnic and sex differences. *British Journal of Psychology*, *70*(3), 351–363.

Jahoda, G. (1992). *Crossroads between culture and mind.* New York: Harvester Wheatsheaf.

Jahoda, G., & McGurk, H. (1974). Pictorial depth perception in Scottish and Ghanaian children. *International Journal of Psychology*, *9*(4), 255–267.

Jahoda, G., & Stacey, B. (1970). Susceptibility to geometrical illusions according to culture and professional training. *Perception and Psychophysics*, *7*(3), 179–184.

Jones, P. A. (1974). Intra-cultural differences in susceptibility to geometrical illusions and pictorial depth perception. *Perceptual and Motor Skills*, *38*, 188–190.

Kilbride, P. L., & Leibowitz, H. W. (1975). Factors affecting the magnitude of the Ponzo perspective illusion among the Baganda. *Perception and Psychophysics*, *17*(6), 543–548.

Kilpatrick, F. P. (1961). *Explorations in transactional psychology.* New York: New York University Press.

Kling, J. W., & Riggs, L. A. (Eds.). (1972). *Woodworth and Schlosberg's experimental psychology.* London: Methuen.

Kunnapas, T. M. (1955). An analysis of the "vertical–horizontal illusion." *Journal of Experimental Psychology*, *49*(2), 134–140.

Leibowitz, H. W., Brislin, R., Perlmutter, L., & Hennessy, R. (1969). Ponzo perspective illusion as a manifestation of space perception. *Science*, *166*, 1174–1176.

Leibowitz, H. W., & Pick, H. A. (1972). Cross-cultural and educational aspects of the Ponzo perspective illusion. *Perception and Psychophysics*, *12*(5), 430–432.

Medvedev, L. I., & Shoshina, I. I. (2004). Quantitative assessment of the effects of hemispheric asymmetry on the distortion of visual perception of the Jastrow version of the Poggendorff figure. *Human Physiology*, *30*(5), 505–510.

Michael, D. N. (1953). A cross-cultural investigation of closure. *Journal of Abnormal and Social Psychology*, *48*(2), 225–230.

Miller, R. J. (1973). Cross-cultural research in perception of pictorial materials. *Psychological Bulletin*, *80*, 135–150.

Morgan, P. (1959). A study in perceptual differences among cultural groups in Southern Africa, using tests of geometric illusions. *Journal of the National Institute of Personnel Research*, *8*, 39–43.

Mundy-Castle, A. C., & Nelson, G. K. (1962). A neuropsychological study of the Knysna forest workers. *Psychologia Africana*, *9*(2), 240–272.

Omari, I. M., & McGinitie, W. H. (1974). Some pictorial artefacts in studies of African children's pictorial depth perception. *Child Development*, *44*, 535–539.

Osgood, C. E. (1962). *Method and theory in experimental psychology.* New York: Oxford University Press.

Peck, L., & Hodges, A. B. (1937). A study of racial differences in eidetic imagery of preschool children. *The Pedagogical Seminary Journal of Genetic Psychology*, *51*, 141–161.

Pedersen, D. M., & Wheeler, J. (1983). The Müller-Lyer illusion among Navajos. *Journal of Social Psychology*, *121*, 3–6.

Piaget, J. (1969). *The mechanism of perception.* London: Routledge & Kegan Paul.

Pollack, R. H. (1965). The detectability threshold as a function of chronological age. *Psychonomic Science*, *3*, 351–352.

Pollack, R. H. (1966). Non-developmental age changes in perception of children. In N. S. Jenkin & R. H. Pollack (Eds.), *Perceptual development: Its relation to theories of intelligence & cognition. Proceedings of a conference sponsored by Institute for Juvenile Research* (pp. 82–102). Chicago: Institute for Juvenile Research.

Pollack, R. H., & Silvar, S. D. (1967). Magnitude of the Müller-Lyer illusion in children as a function of pigmentation of the fundus oculi. *Psychonomic Science*, *8*(2), 83–84.

Pollanac, R. B. (1977). Illusion susceptibility and adaptation to the marine environment: Is carpentered world hypothesis seaworthy? *Journal of Cross-Cultural Psychology*, *8*(4), 425–434.

Rasmjou, S., Hausmann, M., & Güntürkün, O. (1999). Hemispheric dominance and gender in the perception of an illusion. *Neuropsychologia*, *37*(9), 1041–1047.

Reuning, H. (1988). Testing Bushmen in the Central Kalahari. In S. H. Irvine & J. W. Berry (Eds.), *Human abilities in cultural context* (pp. 453–486). New York: Cambridge University Press.

Reuning, H., & Wortley, W. (1973). Psychological studies of the Bushmen. *Psychologia Africana, Monograph Supplement No. 7.*

Rivers, W. H. R. (1901). Vision. In A. C. Haddon (Ed.), *Reports of the Cambridge Anthropological Expedition to Torres Straits* (pp. 1–132). Cambridge, UK: Cambridge University Press.

Rivers, W. H. R. (1903). Observations on the vision of the Uralis and Sholagas. *Madras Government Museum Bulletin*, *5*(1), 3–18.

Rivers, W. H. R. (1905). Observations on the senses of the Todas. *British Journal of Psychology*, *1*(4), 321–396.

Segall, M. H., Campbell, D. T., & Herskovits, M. J. (1963). Cultural differences in the perception of geometric illusions. *Science*, *139*, 769–771.

Segall, M. H., Campbell, D. T., & Herskovits, M. J. (1966). *The influence of culture on visual perception.* Indianapolis: Bobbs-Merrill.

Serpell, R., & Deręgowski, J. B. (1980). The skill of pictorial perception: An interpretation of cross-cultural evidence. *International Journal of Psychology*, *15*, 145–180.

Silvar, S. D., & Pollack, R. W. (1967). Racial differences in pigmentation of the Fundus Oculi. *Psychonomic Science*, *7*(4), 159–160.

Smith, T. (1973). The susceptibility of Xosa groups to a perspective illusion. *Journal of Social Psychology*, *90*(2), 331–332.

Stevens, S. S. (1951). *Handbook of experimental psychology.* New York: Wiley.

Stewart, V. M. (1973). Tests of the "carpentered world" hypothesis by race and environment in America and Zambia. *International Journal of Psychology*, *8*(2), 83–94.

Stewart, V. M. (1974). Cross-cultural test of the "carpentered world" hypothesis using Ames distorted room illusion. *International Journal of Psychology*, *9*(2), 79–89.

Thouless, R. H. (1933). A racial difference in perception. *Journal of Social Psychology*, *4*, 330–339.

Titchener, E. B. (1916). On ethological tests of sensation and perception with special reference to tests of color vision and tactile discrimination described in the reports of the Cambridge Anthropological Expedition to Torres Straits. *Proceedings of the American Philosophical Society*, *55*(3), 204–236.

Vurpillot, E. (1963). *L'organisation perceptive son rôle dans l'evolution des illusions optico-géometriques.* Paris: Vrin.

Wagner, D. A. (1977). Ontogeny of the Ponzo illusion: Effects of age, schooling and environment. *International Journal of Psychology*, *12*(3), 161–176.

Woodworth, R. S. (1910). Racial differences in mental traits. *Science, 31*(788), 171–186.

Woodworth, R. S. (1938). *Experimental psychology.* New York: Henry Holt.

Zusne, L. (1968). Optical illusions: Output of publications. *Perceptual and Motor Skills*, *27*, 175–177.

Chapter 3
Visual Illusion in a Comparative Perspective

Kazuo Fujita, Noriyuki Nakamura, and Sota Watanabe

WHY DO WE STUDY VISUAL ILLUSIONS IN NONHUMAN ANIMALS?

How do nonhuman animals perceive various illusory figures? This question is intriguing in itself for people who are interested in animal behavior and perception. As an owner of several dogs, I am eager to know how they perceive the world surrounding them. Visual illusions are at the core of such a question because the illusory experience is so striking to us. If shown two bones of the same size but different in shape (like Müller-Lyer figures, for example), which would the dog think was larger?

In fact, from a scientific viewpoint, such a question is not simply a matter of interest. The question is important for several reasons. First, like all of our physical characteristics, our visual function is a product of evolution. Knowing how our evolutionary relatives perceive illusory figures may teach us how our way of perceiving the world has been shaped in the history of the paleontological timescale.

Second, relating the illusory perception of a particular species of animals to its ecology may inform us about the adaptive significance of the illusory perception. Our visual system interprets sensory inputs with various constraints. For instance, we find one concave among convexes in the pattern shown in Figure I.3-1. This is because we hypothesize that the light comes from above. Such constraints usually help us to promptly recognize the environment in typical contexts, but they may sometimes fool us if they are used in some atypical contexts, which may result in various visual illusions. Different species are likely to have established different constraints because the information they must draw from the environment depends on their ecology.

Third, studying visual illusions in nonhuman animals also contributes to theories of illusions in humans. If illusions are a product of evolution, those in humans and nonhumans should be accounted for by similar factors of different contributing weights rather than by completely different sets of factors or mechanisms. Understanding the nature of illusory perception in nonhumans may tell us which factors are essential and which are auxiliary.

Fourth, if illusory perception depends on early visual experience, controlling environment during upbringing may be one way to detect the contribution of experience. It is known that early visual experience can have profound effects on later visual functions (e.g., Held & Hein, 1963; Torii & Mochizuki, 1995). Studies have found variation in the magnitude of illusions (e.g., Jahoda & Stacy, 1970), as well as to what aspects people attend (e.g., Kitayama, Duffy, Kawamura, & Larsen, 2003), by culture and experience.

Despite such importance, we still do not have a solid knowledge of how different species perceive various illusory figures; there are so many species and so many illusory figures, and it takes so much time to demonstrate clearly even just one visual illusion in nonhuman animals. In this chapter we present a brief review of the literature then look at more recent, systematic analyses before focusing on pigeons, the best-studied species regarding nonhuman illusory perception. Last we present a view of human perceptual functions situated in the animal kingdom.

VISUAL ILLUSIONS IN NONHUMAN SPECIES: A BRIEF LOOK AT EARLY STUDIES

The question of whether nonhuman animals experience visual illusion has been asked since the first half of the 20th century. There have been several demonstrations of various illusions in animals; birds were the primary subjects in these early studies. For instance, Révész (1924) tested domestic chickens (hens) for their perception of Jastrow figures, in which two curved bands of the same length are placed inside and outside like a bunch of bananas; the one inside looks larger to human eyes (Robinson, 1972, Pattern 2.75). Seven hens were first trained to choose the smaller of two figures of the same shape, including circles, squares, rectangles, and triangles. When two successful birds were tested with two curved bands of the same size, they tended to choose the one outside. Révész (1924) also tested a bird in a version of the vertical–horizontal illusion of surface size. When two rectangles of the similar sizes were placed horizontally and vertically, the hen tended to choose the horizontally placed figure. But this appears inconclusive, because the clue for the hen to choose from the two figures was not identified; the birds might have simply reacted to the vertical length.

According to Dominguez (1954), Winslow (1933) also reported that domestic chickens perceive the vertical–horizontal illusion (the Fick illusion; Robinson, 1972,

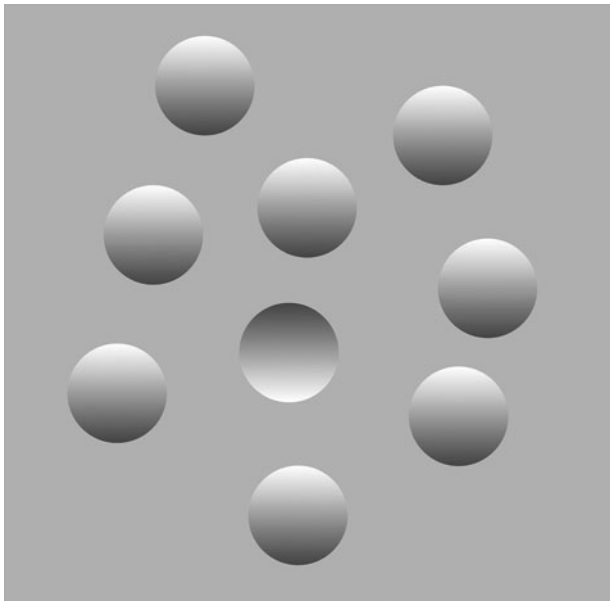


Figure I.3-1. Light comes above. (Original drawing by K. Fujita.)

Pattern 3.76), illusion of interrupted extent (the Oppel-Kundt illusion; Robinson, 1972, Pattern 2.69), the Müller-Lyer illusion (Robinson, 1972, Pattern 2.1), and the breadth of rectangles illusion, in which narrow rectangles are perceived higher than wide rectangles.

Warden and Baar (1929) tested perception of the Müller-Lyer illusion in two ring doves. The doves were trained to choose the shorter of two horizontal target lines with two short vertical lines on both ends (like a flattened H). After transfer of this performance to the lines with two brackets of the same orientation, the birds were tested with Müller-Lyer figures. The birds tended to choose stimuli with outward- over inward-pointing arrowheads, which was consistent with the typical Müller-Lyer illusion. However, the demonstration seems unreliable, as the accuracy of the birds' choice in the last phase before the test remained below 70% and 60% for respective birds.

Dücker (1966, cited in Suzuki, 1995) tested several visual illusions in various species of fish, avians, and mammals. He reported that crucian carp, a chick, avadavats, starlings, a mistle thrush, and guinea pigs were susceptible to the Zöllner illusion (Robinson, 1972, Pattern 3.1); crucian carp, avadavats, a starling, and guinea pigs to the Ebbinghaus–Titchener circles illusion (Robinson, 1972, Pattern 2.63); and crucian carp, a starling, and an avadavat to the Orbison illusions, in which a square placed on the concentric arcs looks curved (Robinson, 1972, Pattern 3.20).

Malott, Malott, and Pokrzywinski, (1967) tested the Müller-Lyer illusion in pigeons. They trained pigeons to peck at a flattened H-shaped figure of a particular length on a translucent key. Then the target lines of varied lengths with inward- or outward-pointing brackets were presented on the key. The peak frequency of responses by birds was shifted toward longer for the lines with outward-pointing brackets, which illustrates underestimation of the target line length. Unfortunately, no shift showing overestimation was found for the lines with inward-pointing brackets. Thus the demonstration was inconclusive.

Studies on illusory perception in nonhuman primates started in the second half of the 20th century. Dominguez (1954) tested three species of monkeys (two rhesus macaques, a sooty mangabey, and a white-throated capuchin monkey) for their perception of the horizontal–vertical illusion of surface. The monkeys were trained to choose a small square of a particular surface size among the two constructed with vertical or horizontal grids (the Helmholz's squares; Robinson, 1972, Pattern 2.73). Then they were shown the two squares of the same size, one constructed with vertical grids and the other with horizontal grids. The monkeys tended to choose the square of the vertical grids, which suggested that they perceived it as larger than the square of the horizontal grids. This may not be conclusive, however, because there is no guarantee that the monkeys responded on the basis of the surface size. Dominguez also obtained positive evidence for horizontal–vertical illusion of the length (the Fick illusion; Robinson, 1972, Pattern 3.76) and the breadth of rectangles illusion, in which narrower rectangles look higher than wider ones of the same height using a similar procedure.

Harris (1968, cited in Davis, 1974) tested four stumptailed macaques for the perception of the breadth of rectangles illusion. In correcting the weakness in the Dominguez (1954) study, Harris changed the procedure into a decision of relative height; monkeys received an occasional reward by choosing the higher of the two rectangles varying in size. In a later test with the rectangles of the same height but different width, the monkeys tended to choose the figure that was narrower in width.

Behnar and Samuel (1982) tested two anubis baboons for their perception of the Zöllner illusion. The baboons were shown trios of cards on which four vertical lines were drawn and were trained to choose the odd one regarding the orientations of the lines; for instance, the baboons were given a choice among the one that consisted of four parallel lines and two that consisted of four alternatingly converging lines (Fig. I.3-2, top). Then they were tested with the options including a Zöllner figure like the ones shown in

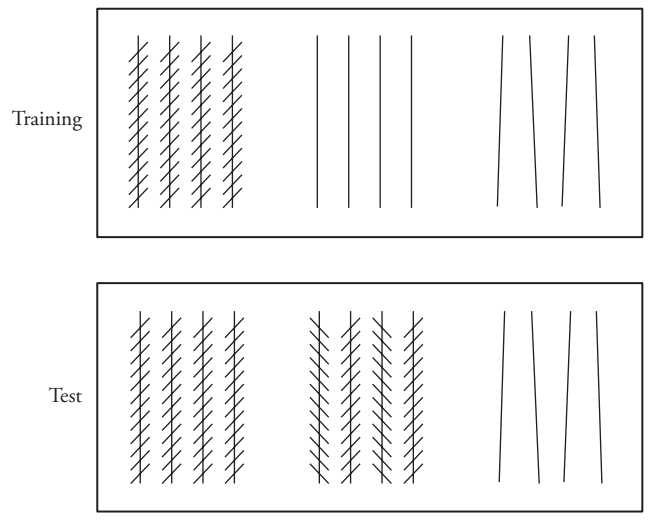


Figure I.3-2. Training (top) and test figures (bottom) to demonstrate Zöllner illusion in baboons. (Adapted from Benhar & Samuel, 1982.)

the bottom of Figure I.3-2. As a result, the baboons tended to choose the leftmost pattern, which suggested that they perceived the Zöllner figure as alternatingly converging and diverging. However, because of the reason described later, it may not mean that baboons perceived the Zöllner figures as humans do.

Bayne and Davis (1983) showed that rhesus macaques perceive a version of the Ponzo illusion (Robinson, 1972, Pattern 3.32a). The monkeys were trained to respond on one of the locations for the rectangle sample and the other for the trapezoid of the longer upside. Then the monkeys were tested with the rectangle superimposed on several upward converging lines. The subjects responded on the location corresponding to the trapezoid.

As we can see, these early studies in general showed that various nonhuman animals, including fish, birds, rodents, and primates, also perceive many kinds of geometric illusions as humans do. However, most of the demonstrations remain superficial, and it is difficult to conclude whether the illusory perception of nonhumans is homologous to that of humans. Systematic analyses and comparison of the nature of illusory perception under various parametric conditions are needed.

MORE RECENT STUDIES AND SPECIES DIFFERENCES DETECTED IN ILLUSORY PERCEPTION

Ponzo Illusion

In a series of studies, Fujita and colleagues tested pigeons and nonhuman primates for their perception of the Ponzo illusion in various parametric conditions (Fujita, 1996. 1997; Fujita, Blough, & Blough, 1991; 1993). The general procedure was to train subjects to classify horizontal bars of six different lengths into "long" and "short" by differentially touching two response locations on the touch-sensitive monitor. Only a single bar was presented at a time. The initial training was done with isolated bars, then the illusion-inducing context lines—converging lines in the case of typical Ponzo figures—were superimposed at a fixed position (Figure I.3-3a, middle context). After learning this discrimination, the subjects were tested with the figures having the context lines either above or below the standard position. If the subjects perceive the Ponzo illusion like humans, the classification of "long" and "short" should be biased toward "long" for the figures having the apex of the converging context closer to the target bar (Fig. I.3-3a, low context) and toward "short" for those having the apex farther from the target (Figure I.3-3a, high context). The subjects received a reward (or a conditional reward) irrespective of the choice of the response keys for these test figures.

Figure I.3-4 shows the results for the figures with the most basic inverted-V context (two upward-converging lines) in pigeons, rhesus macaques, and a chimpanzee (Fujita, 1997; Fujita, et al., 1993). All three species tested showed a predicted bias of choices; the proportion of the choice for "long" was more frequent for the low-context figures (stars) than for high-context figures (squares). Therefore the three species tested commonly perceived a version of the Ponzo illusion.

A possible cause of the Ponzo illusion is the perspective impression of the converging context lines. That is, the bar located near the apex of the converging lines is perceived as more distant than the one located farther from the apex, and our perceptual system applies a size constancy scaling to automatically correct the subjective size impression (misapplied constancy scaling hypothesis; Gregory, 1963).

In order to test this popular hypothesis, we compared two types of contexts, labeled perspective and nonperspective contexts. The former consisted of eight lines converging toward a focus (Fig. I.3-3b), and the latter had four parallel lines on each side (Fig. I.3-3c). The innermost lines were identical to the inverted-V context and common to the two types of context. However, no species, including pigeons, rhesus macaques, chimpanzees, and humans, showed more enhanced illusion for the perspective context. In fact, the magnitude of the illusion was almost the same among the three contexts: inverted-V, perspective, and nonperspective (Fujita, 1997; Fujita, et al., 1991, 1993). Thus, for these four species, the Ponzo illusion of this version is mainly determined by the innermost lines; stronger perspective implied by the multiple converging lines has little contribution.

However, when we replaced the context from lines to short vertical lines having the same gap against the target bar as the line contexts (Fig. I.3-3d, dot context), we found a large species difference among the three species of primates. In humans, the magnitude of the illusion was doubly large for the dot context compared with the inverted V. On the other hand, rhesus macaques showed no illusion effect for the dot context. Chimpanzees showed an effect comparable to that of the inverted-V context. Thus there was a trend from monkeys to humans according to the phylogenetic distance (Fujita, 1997).

Another difference was found between rhesus macaques and humans when we tested figures superimposed on perspective photos. There were four types of stimuli, shown in Figure I.3-5. The first was the inverted-V context superimposed on the upright photo (upright photo + inverted V, first [leftmost] column); the second was the inverted-V context on the inverted photo (inverted photo + inverted V, second

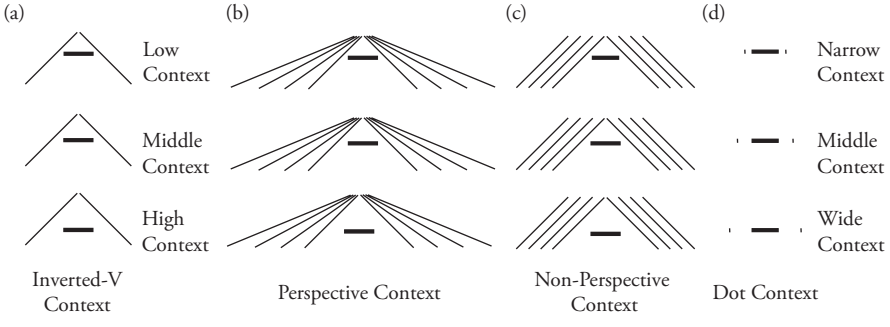


Figure I.3-3. Stimuli used by a series of studies of the Ponzo illusion in pigeons and nonhuman primates by Fujita and colleagues. (Adapted from Fujita, 1997; Fujita, Blough, & Blough, 1991, 1993.)

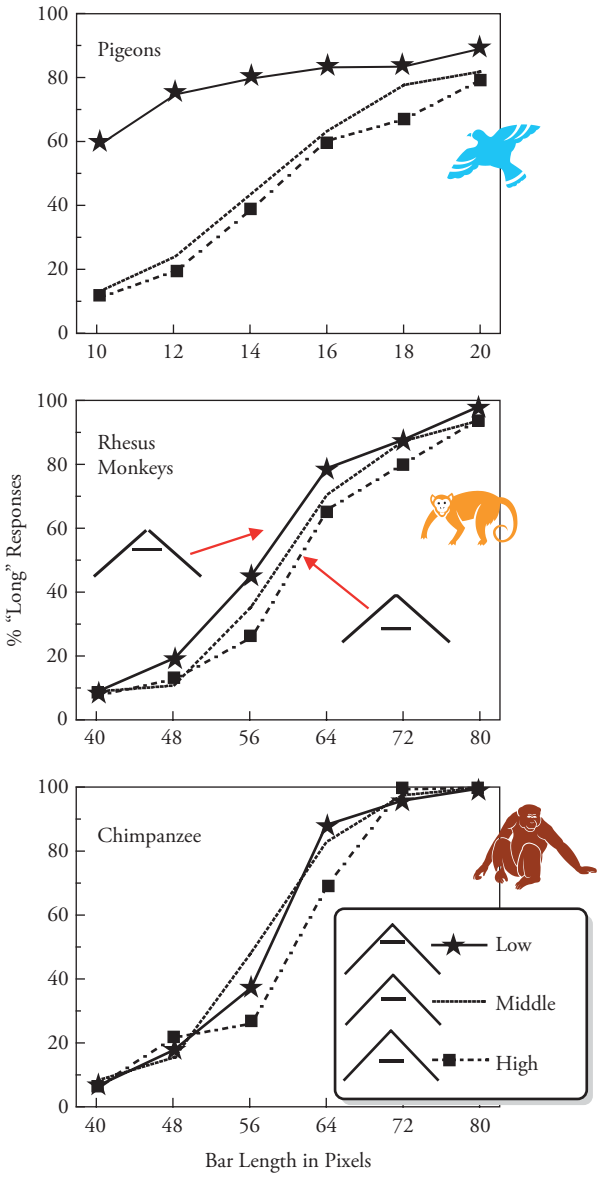


Figure I.3-4. Ponzo illusion in pigeons (top), rhesus macaques (middle), and a chimpanzee (bottom). Horizontal axis shows the length of target bar and the vertical axis shows proportion of responses for "long." The proportion of responses for "long" is more frequent for the low-context condition (stars) than for the high-context condition (squares). Dotted lines with no symbols are for the middle-context condition. (Adapted from Fujita, 1997; Fujita et al., 1993.)

column from left); the third was no line context on the upright photo (upright photo only, third column); and the last was no line context on the inverted photo (inverted photo only, rightmost column). Humans perceived illusion for the two types of stimuli having upright photos (first and third columns). Almost no illusion occurred for the other two stimuli having inverted photos. On the other hand, rhesus macaques showed an illusory effect for the stimuli having an inverted-V context irrespective of the orientation of the photo (left two columns). However, they also perceived a small but significant illusion for the upright photo with no line context (third column). This means that, unlike the converging lines, the photographic perspective impression has a significant effect on inducing Ponzo-like illusions for both species. But the relative strength of the context lines and that of the photographic perspective differs between species; the former is stronger for rhesus macaques, and it overshadows any effects of the latter, whereas the reverse is true for humans (Fujita, 1996).

The effect of photographic perspective to induce illusion in the target size has been evidenced in horses too (Timney & Keil, 1996). Also, when the line-drawn perspective is strong enough, nonhuman primates perceive an overestimation of the target stimuli placed "farther" compared with those placed "closer" (the corridor illusion; Guinea baboons: Barbet & Fagot, 2002; chimpanzees: Imura, Tomonaga, & Yagi, 2008). Many species may have sensitivity to pictorial depth in perception of illusion induced by the misapplied constancy scaling (Gregory, 1963). Probably the strength of this effect can greatly vary among species.

Müller-Lyer Illusion

We tested pigeons for their perception of the Müller-Lyer illusion using a similar task (Nakamura, Fujita, Ushitani, & Miyata, 2006). Pigeons were first trained to classify the horizontal lines into "long" and "short." Then, after being accustomed to the brackets of the same orientation at both ends of the target lines (Figure 6a, top), the birds were tested with the Müller-Lyer figures (Fig. I.3-6a, middle and bottom). Their classification responses on the test trials were nondifferentially reinforced as in the Ponzo studies.

The pigeons showed a consistent tendency to classify the lines with inward-pointing brackets (triangles in Fig. I.3-7) into "long" more often than those with outward-pointing brackets (squares). Careful analyses confirmed that other potential clues such as the entire length of the figures and

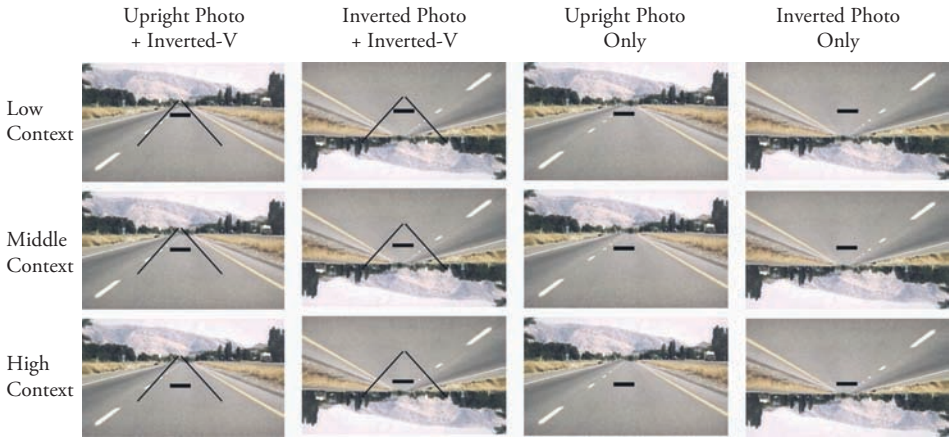


Figure I.3-5. Stimuli used in the test of photographic perspective on the Ponzo-like illusion in rhesus macaques. (Adapted from Fujita, 1996.)

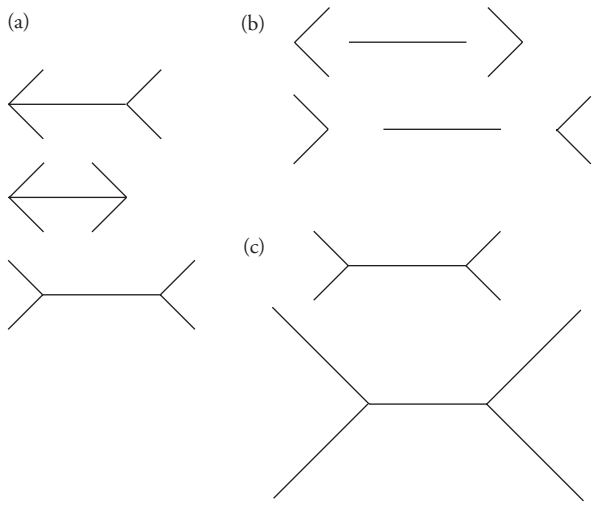


Figure I.3-6. Variations of the Müller-Lyer figures. (Original drawing by K. Fujita.)

the width of the gap of the two brackets could not account for the data obtained. The results strongly suggest that pigeons perceive the Müller-Lyer illusion (Nakamura et al., 2006).

Is the illusory perception of Müller-Lyer figures by pigeons based on the same mechanism as that by humans? We asked this question by using a few parametric changes in the figures. Humans are known to perceive a reversed Müller-Lyer illusion when the brackets are detached from the ends of the target lines (Fig. I.3-6b; Yanagisawa, 1939). However, we failed to find this effect in pigeons whereas we confirmed it in humans using exactly the same stimuli used for pigeons (Nakamura et al., 2006).

Another reversed illusion is known in humans. When the inward-pointing brackets are extended, the target line starts to shrink (Fig. I.3-6c; Obonai, 1954). We tested this in pigeons but again failed to obtain the effects, in contrast to the effect in humans tested with the same figures (Nakamura, Watanabe, & Fujita, 2009).

What makes such a difference between humans and pigeons in the perception of Müller-Lyer and derivative figures? We suspect that pigeons might be susceptible to

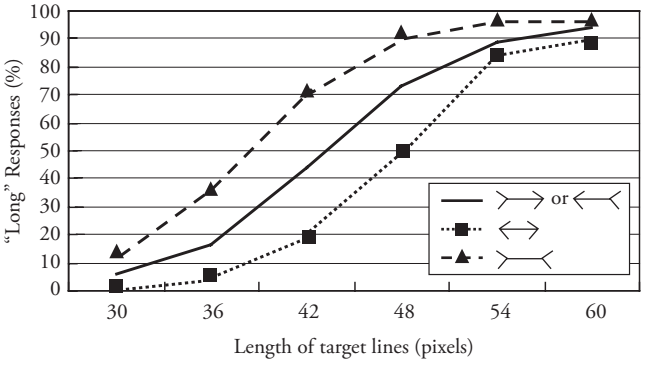


Figure I.3-7. Müller-Lyer illusion in pigeons (average for three birds). The proportion of responses for "long" is more frequent for the figures with inward-pointing brackets (triangles) than for those with outward-pointing brackets (squares). The solid line with no symbols is for the baseline Judd figures. (Adapted from Nakamura et al., 2006.)

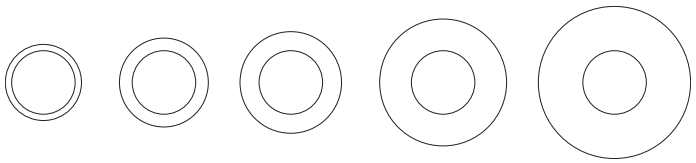


Figure I.3-8. Examples of how assimilation and contrast effects work. (Original drawing by K. Fujita.)

assimilation effects but not to *contrast* effects. *Assimilation* and *contrast* are two fundamental functions that would work in various illusory figures. Goto et al. (2007) describes *assimilation* and *contrast* as "a perceived distortion among area, size, length, angle, and curvature in adjacent stimulus configurations based on unification (approach; decrease of difference) and distinction (separation; increase in difference)" (p. 42). Assimilation is more likely to occur between figures placed close to each other whereas contrast is more likely to occur between figures placed somewhat more distant to each other. The drawing in Figure I.3-8 (concentric circles illusion) clearly demonstrates a gradual change of the effect from assimilation (leftmost figure) to contrast (rightmost figure). Note that all inner circles are the same in their actual size.

The two types of reversed Müller-Lyer illusions in humans are a consequence of the *contrast* effect because the effects come from more distantly placed parts of the figures than the standard Müller-Lyer figures, for which the *assimilation* effect is more likely to cause the illusion. If this is true, pigeons may not see the typical illusory figures for which we observe a striking contrast effect, like Ebbinghaus–Titchener circles illusion, as humans see them.

Ebbinghaus–Titchener Circles Illusion

We first trained pigeons to classify the size of the black disks into "small" and "large" by training them to peck at two response locations (Nakamura, Watanabe, & Fujita, 2008), just like in the previous Ponzo and Müller-Lyer studies. Next, we gradually faded in six disks from light grey to black surrounding the target disk. The size of these peripheral disks was constant (Fig. I.3-9, top). After the pigeons learned to correctly classify the target disks by ignoring the peripheral ones, we presented the target disks with small and large peripheral ones, just like the typical Ebbinghaus–Titchener circles figure (Fig. I.3-9, bottom).

Figure I.3-10 shows the results. All five pigeons tested showed a consistent shift in the proportion of the choice for "large" for test figures compared with training figures. Thus they perceived illusion in size, but the shift was in the direction opposite to what is expected by the typical Ebbinghaus–Titchener circles illusion in humans. That is, the birds estimated the size of the disk surrounded by large disks to be larger and that surrounded by small disks to be smaller than that of training stimulus (Nakamura, et al., 2008).

The pigeons could not have responded to the size of surrounding disks because, if so, the proportion of "large" responses should have been flat irrespective of the size of the target disk. Might the pigeons have used weighted average

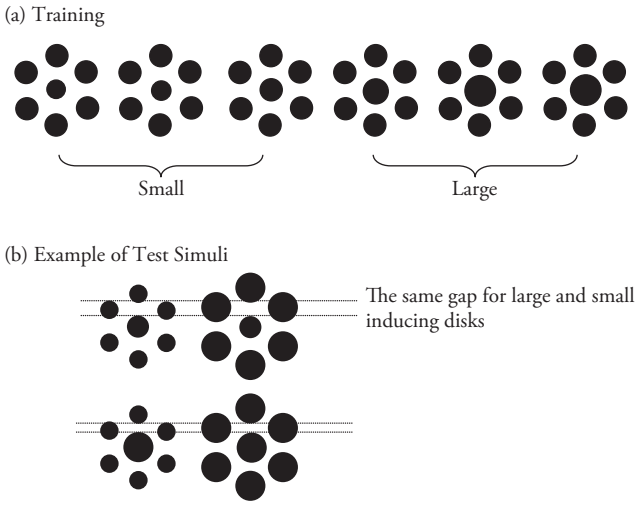


Figure I.3-9. Stimuli used to test perception of the Ebbinghaus–Titchener circles illusion in pigeons. (a) training figures and (b) examples of test figures. (Adapted from Nakamura et al., 2008.)

of all the disks in the display? To examine this possibility, we tested the pigeons with isolated target disks differing in size (Nakamura et al., 2008). If the pigeons had responded on the weighted average, the proportion of "large" responses should have changed more sharply as a function of the size of the target disk for the isolated target disk than for the target surrounded by other disks. The result clearly showed that this was not the case; the curves of the proportion of "large" were almost exactly the same. These results suggest that pigeons in fact perceive the reversed Ebbinghaus–Titchener circles illusion. That means pigeons experience assimilation, rather than contrast, of the target disk to surrounding disks.

This appears almost unbelievable to the human eye, but a recent study by Parron and Fagot (2007) shows that Guinea baboons fail to perceive the Ebbinghaus–Titchener circles illusion. Also, it is known that humans perceive overestimation of the target disk when the farther portion of the surrounding disks is erased (Morinaga, 1956; Fig. I.3-11). In this case, the target disk assimilates to surrounding contours.

Zöllner Illusion

We recently tested pigeons for their perception of Zöllner figures (Watanabe, Nakamura, & Fujita, 2011). We presented a pair of lines of clearly different inclinations and trained the birds to peck at the narrower (or wider, for some birds) of the two gaps at the top and the bottom ends of the lines (see Fig. I.3-12a for an example of training stimuli). The crosshatches on the lines appeared in random orientation across trials. Then we tested the pigeons with the Zöllner-like patterns with varied convergences of lines, including parallel ones (Fig. I.3-12b).

Due to the orientation of the crosshatches, there were two types of test stimuli, labeled Boost-upside-wider and Boost-downside-wider patterns, respectively. These labels come from human subjective impression.

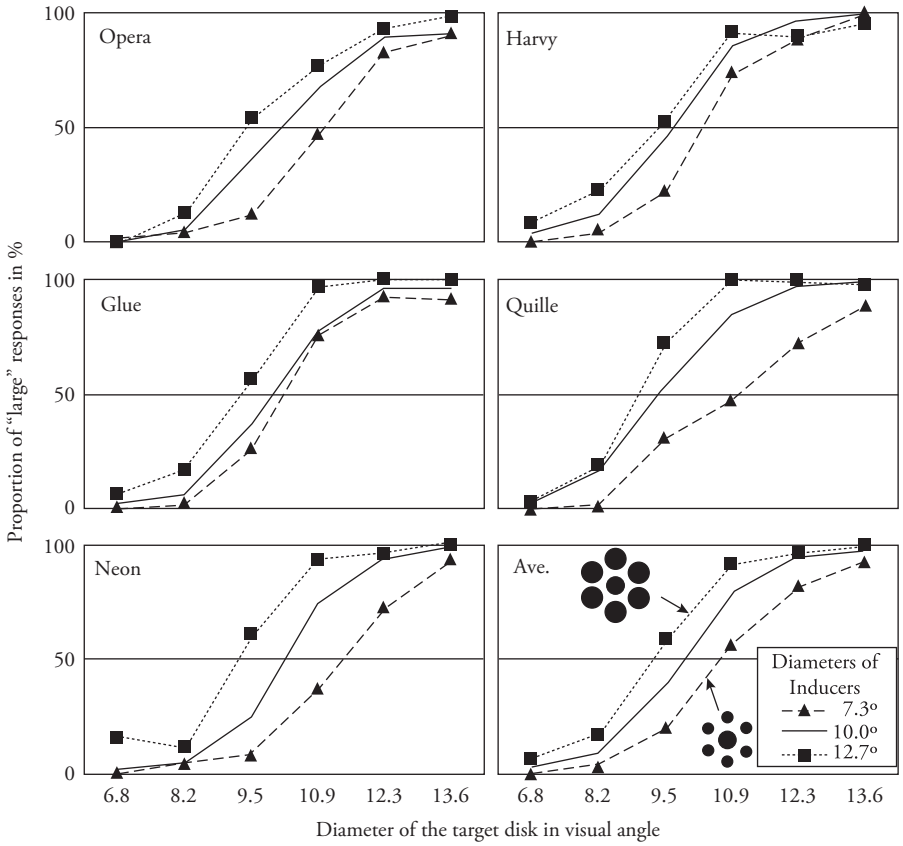


Figure I.3-10. Reversed Ebbinghaus–Titchener circles illusion in individual pigeons. The bottom right graph is for the average. All five birds responded "large" more often for the target disk surrounded by large disks (squares) than that surrounded by small disks (triangles). The solid line with no symbols is for the baseline with middle-sized surrounding disks. (Adapted from Nakamura et al., 2008.)

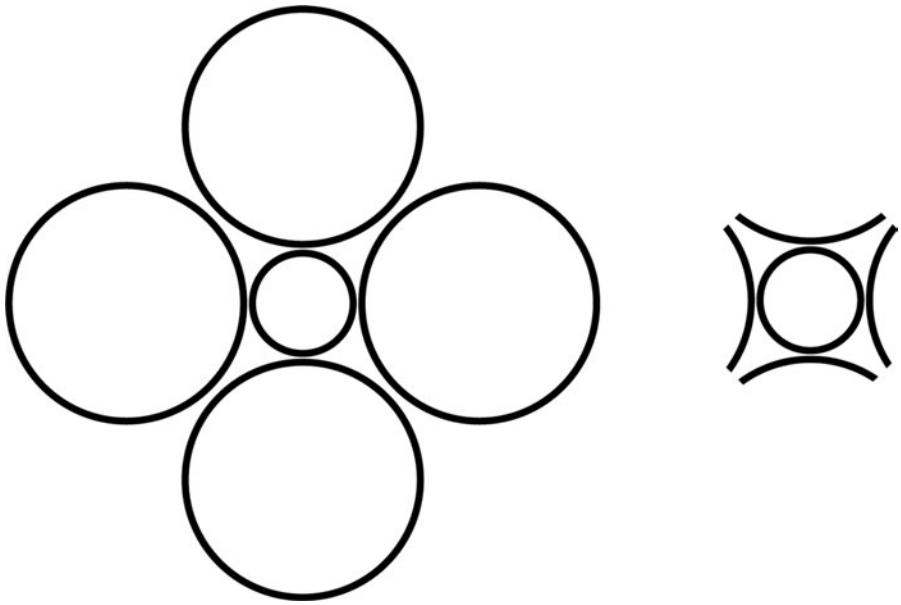


Figure I.3-11. A reversed Ebbinghaus–Titchener circles illusion for humans. (Adapted from Morinaga, 1956.)

Figure I.3-13 shows the proportion of responses showing the report that "the top is wider" as a function of the convergence of the lines. The top panel is for pigeons, and the bottom panel is for humans who did the same task. Both species shifted their responses for the two types of test figures (squares for Boost-downside-wider and triangles for Boost-upside-wider). But the shift occurred in the opposite directions between pigeons and humans. In other words, pigeons do perceive the illusory change in the orientation of the main lines but in the direction opposite to humans.

Although this distortion is in the orientation dimension, the same argument of pigeons' assimilation dominance may again apply here. That is, pigeons may underestimate the difference in orientation of two lines at the crossing. In other words, the orientation of the lines may assimilate to each other. Thus the main line embedded in many crosshatches may be perceived as slanted toward the orientation of the crosshatches. Let us take the example of Boost-upside-wider pattern in Figure I.3-12b. The main line at the top may be perceived as slanted toward the horizontal (the crosshatch orientation) and the main line at the bottom, on the contrary, slanted toward the vertical. Hence the two lines are perceived as converging upward, resulting in a reversed Zöllner illusion. In contrast, humans overestimate acute angles (Kitaoka & Ishihara, 2000; Oyama, 1975), which means the distortion of the angles is in the direction opposite to that in pigeons. It may be noted that humans also perceive a reversed Zöllner illusion when the crosshatches make a very small (c. 3°) angles

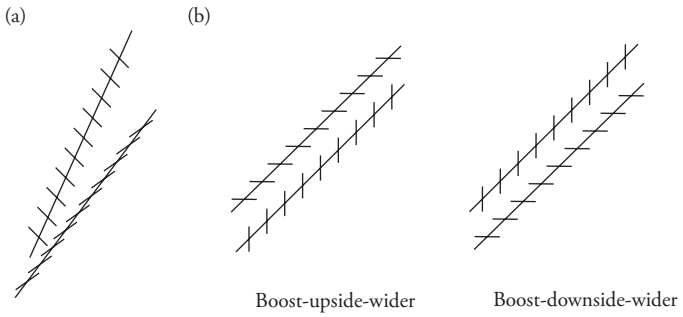


Figure I.3-12. Stimuli used to test perception of Zöllner figures in pigeons. (a) an example of training figures and (b) examples of test figures. (Adapted from Watanabe, Nakamura, & Fujita, 2011.)

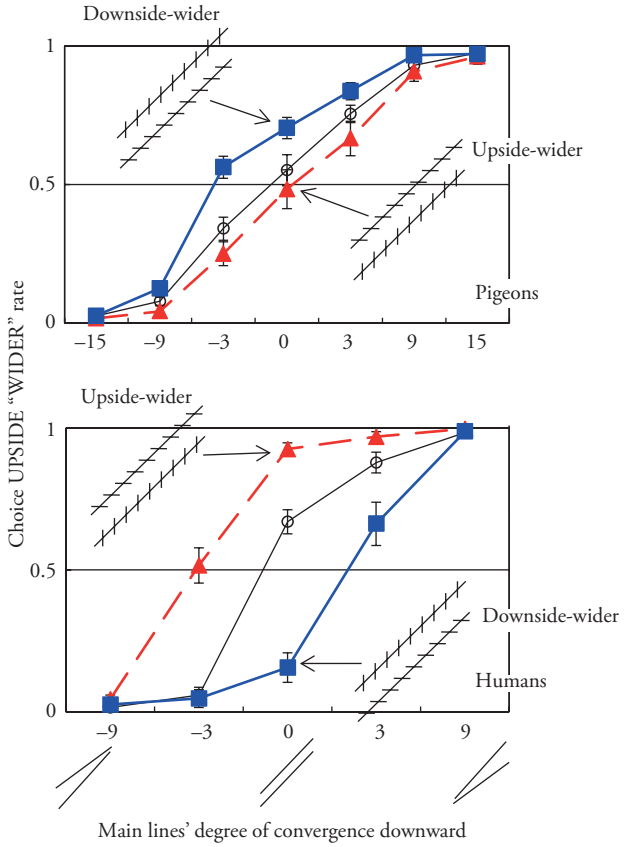


Figure I.3-13. Reversed Zöllner illusion in pigeons (top) and the typical Zöllner illusion in humans tested in the same task. The increase/decrease in the proportion of responses implicating two target lines diverging upward is completely reversed for pigeons and humans. (Adapted from Watanabe et al., 2011.)

against the main lines (Imai, 1962), as these small angles are underestimated.

DISCUSSION: DIVERSITY FOUND IN THE PERCEPTUAL WORLD

Although earlier studies have suggested that nonhuman animals perceive various illusions similar to the way humans do, more recent studies have revealed that there may be great diversity in how species perceive various illusory figures. We described several cases showing this point. Rhesus macaques, chimpanzees, and humans similarly perceive the standard Ponzo illusion, but they perceive a version of this with short lines replaced the converging lines (Fig. I.3-3d, dot context) differently; the illusory effect is stronger for the dot context than for the regular inverted-V context in humans, comparable in chimpanzees, and no illusion for the former in rhesus macaques. Although the perspective photo has an effect in inducing Ponzo-like illusions in rhesus macaques and humans, this effect is weaker than the line context embedded in the photo in the former species, whereas the reverse is true in the latter. Thus even among closely related primates, illusory perception may be different.

This difference among primates might still be viewed as quantitative. But comparison of birds and humans has suggested that there may be a difference in quality. Not only do pigeons fail to experience versions of reversed Müller-Lyer illusions, at least in two cases, pigeons also perceive illusions in the direction opposite to how humans do. For Ebbinghaus–Titchener figures, pigeons perceive overestimation of the size of the target disk surrounded by large disks and underestimation of it surrounded by small disks. They also perceive the orientation of subjective convergence in the Zöllner figures opposite to how humans do. As we discussed earlier, this difference may be accounted for by the tendency of pigeons to assimilate the length, size, and orientations of the figures placed close to each other.

We recently tested bantam chickens for how they perceive Ebbinghaus–Titchener (Nakamura, Watanabe, & Fujita, 2014) and Zöllner figures (Watanabe, Nakamura, & Fujita, 2013). In fact, the results were the same as in pigeons. Thus this surprising perception might be shared in some of the avian species. More species have to be tested in the future.

Pigeons have been shown to have a very locally oriented vision. For instance, unlike humans, pigeons are better at detecting local difference than global difference in hierarchically organized stimuli, such as a large "T" made of small "Os" (the local precedence effect: Cavoto & Cook, 2001). When pigeons judge whether two dots are in the same visual object, they tend to use clues in the vicinity of the dots and not the whole object (Lazareva, Levin, Vecera, & Wasserman, 2006). Pigeons perceive cartoons by decomposing them into parts, rather than attending to the global arrangements, and scrambling the arrangement barely disrupts discrimination (Cerella, 1980; Watanabe, 2001).

Pigeons have been repeatedly shown not to amodally complete partly occluded figures (Aust & Huber, 2006; Cerella, 1980; Fujita, 2001; Sekuler, Lee, & Shettleworth, 1996), even if a pair of visible portions have a common motion (Ushitani, Fujita, & Yamanaka, 2001), for which even four-month-old infants perceive unity of the two (Kellman & Spelke, 1983). In contrast, nonhuman primate species so easily show amodal completion in various tasks (e.g., Deruelle, Barbet, Dépy, & Fagot, 2000; Fujita & Giersch, 2005; Nagasaka & Osada, 2000; Sato, Kanazawa, & Fujita, 1997; Sugita, 1999), and some other mammals (Kanizsa, Renzi, Conte, Compostela, & Guerani, 1993), birds (Lea, Slater, & Ryan, 1996; Regolin & Vallortigara, 1995; Tvardíková & Fuchs, 2010, but see Nakamura, Watanabe, Betsuyaku, & Fujita, 2010; Nakamura et al., 2011, for the contradictory data in chickens), and fish (Sovrano & Bisazza, 2008) may have been suggested to do so. In view of this, such failure by pigeons to complete appears peculiar and even maladaptive. However, this failure could sometimes merit them; they meet no difficulty in detecting a punched figure when the punched portion is exactly covered by another figure (Fujita & Ushitani, 2005), which causes serious difficulty in humans because they automatically complete the occluded contour (Rauschenberger & Yantis, 2001). Thus, the local- rather than global-oriented vision in pigeons can be adaptive to the life of pigeons, whose primary agenda in feeding is to pick up small grains as they find the seeds.

Now let us go back to the points raised at the introduction. We are able to provide some answers and comments on the first three points.

The first question was the evolutionary history. We have demonstrated that visual systems that may experience illusion in some cases may have a long history of evolution. There remains the possibility that birds and primates have independently developed a converged system, but this possibility may not be so likely, as earlier studies demonstrated some fish and nonprimate mammals experience various illusions. More species in wider taxa must be tested for their perception of various visual illusions in order to draw a better sketch of the evolution of illusory perception.

The second question was the relationship between illusion and ecology. In this aspect, pigeons are a good case in point; although they do experience illusion like humans, they sometimes perceive reversed illusions. As we discussed, this difference may come from a very local-oriented visual processing in pigeons also suggested by various aspects other than geometric illusion. We have suggested that this may be related to their feeding ecology, but this hypothesis must be tested in other grain-eating species.

Regarding the third point, we have shown that relative weights on assimilation and contrast may account for striking a difference in some visual illusions, which was demonstrated in reversed Ebbinghaus–Titchener circles and Zöllner illusions. We also provided evidence against the misapplied constancy scaling hypothesis as an account for the Ponzo illusion.

Finally, we would like to stress that, by analyzing perception of various illusory figures in various nonhuman species, we have found astonishing diversity in visual processing among species. This is not because they have different peripheral visual systems; both birds and nonhuman primates have fine vision at the receptor level, and the primary information is conveyed to and processed in their central nervous system. Thus the difference lies in the evolutionary tactic of what they extract from the sensory inputs

and how they deal with the extracts. Because the needs are different from species to species, this tactic should vary correspondingly. In view of this, human perception is also a product of this evolutionary tactic that matches to human needs. It is not an ultimate solution in the animal kingdom. In fact it is no more than one of various solutions installed in diverse species.

ACKNOWLEDGMENTS

Preparation of this chapter was supported by the Grants-in-Aid for Scientific Research (S), No.20220004, and (A), 25240020, from the Japan Society for the Promotion of Sciences to KF.

REFERENCES

Aust, U., & Huber, L. (2006). Does the use of natural stimuli facilitate amodal completion in pigeons? *Perception*, *35*, 333–349.

Barbet, I., & Fagot, J. (2002). Perception of the corridor illusion by baboons (*Papio papio*). *Behavioural Brain Research*, *132*, 111–115.

Bayne, K. A. L., & Davis, R. T. (1983). Susceptibility of rhesus monkeys (*Macaca mulatta*) to the Ponzo illusion. *Bulletin of the Psychonomic Society*, *21*, 476–478.

Benhar, E., & Samuel, D. (1982). Visual illusions in the baboon (*Papio anubis*). *Animal Leaning & Behavior*, *10*, 113–118.

Cavoto, K. K., & Cook, R. G. (2001). Cognitive precedence for local information in hierarchical stimulus processing by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *27*, 3–16.

Cerella, J. (1980). The pigeons' analysis of pictures. *Pattern Recognitio*n, *12*, 1–6.

Davis, R. T. (1974). Monkeys as perceivers. In L. A. Rosenblum (Ed.), *Primate behavior* (Vol. 3, pp. 115–142). New York: Academic Press.

Deruelle, C., Barbet, I., Dépy, D., & Fagot, J. (2000). Perception of partly occluded figures by baboons (*Papio papio*). *Perception*, *29*, 1483–1497.

Dominguez, K. E. (1954). A study of visual illusions in the monkey. *The Journal of Genetic Psychology*, *85*, 105–127.

Dücker, G. (1966). Untersuchungen über geometrisch-optische Tauschungen bei Wiebeltieren. *Zeitschrift für Tierpsychologie*, *23*, 452–496.

Fujita, K. (1996). Linear perspective and the Ponzo illusion: A comparison between rhesus monkeys and humans. *Japanese Psychological Research*, *38*, 136–145.

Fujita, K. (1997). Perception of the Ponzo illusion by rhesus monkeys, chimpanzees, and humans: Similarity and difference in the three primate species. *Perception & Psychophysics*, *59*, 284–292.

Fujita, K. (2001). Perceptual completion in rhesus monkeys (*Macaca mulatta*) and pigeons (*Columba livia*). *Perception & Psychophysics*, *63*, 115–125.

Fujita, K., Blough, D. S., & Blough, P. M. (1991). Pigeons see the Ponzo illusion. *Animal Learning & Behavior*, *19*, 283–293.

Fujita, K., Blough, D. S., & Blough, P. M. (1993). Effects of the inclination of context lines on perception of the Ponzo illusion by pigeons. *Animal Learning & Behavior*, *21*, 29–34.

Fujita, K., & Giersch, A. (2005). What perceptual rules do capuchin monkeys (*Cebus apella*) follow in completing partly occluded figures? *Journal of Experimental Psychology: Animal Behavior Processes*, *31*(4), 387–398.

Fujita, K., & Ushitani, T. (2005). Better living by not completing: A wonderful peculiarity of pigeon vision? *Behavioural Processes*, *69*, 59–66.

Goto, T., Uchiyama, I., Imai, A., Takahashi, S., Hanari, T., Nakamura, S., & Kobari, H. (2007). Assimilation and contrast in optic illusions. *Japanese Psychological Research*, *49*, 33–44.

Gregory, R. L. (1963). Distortion of visual space as inappropriate constancy scaling. *Nature*, *199*, 678–680.

Harris, A. V. (1968). Perception of the horizontal-vertical illusion by stumptail monkeys. *Radford Review*, *22*, 61–72.

Held, R., & Hein, A. (1963). Movement-produced stimulation in the development of visually guided behavior. *Journal of Comparative and Physiological Psychology*, *56*, 872–876.

Imai, S. (1962). Zöllner gyaku-sakushi ni kansuru kenkyu [A study on a reversed Zöllner illusion]. *Tokyo Metropolitan University Jinmon Gakuho*, *30*, 31–44. (in Japanese)

Imura, T., Tomonaga, M., & Yagi, T. (2008). The effects of linear perspective on relative size discrimination in chimpanzees (*Pan troglodytes*) and humans (*Homo sapiens*). *Behavioural Processes*, *77*, 306–312.

Jahoda, G., & Stacy, B. (1970). Susceptibility to geometrical illusions according to culture and professional training. *Perception & Psychophysics*, 7, 179–184.

Kanizsa, G., Renzi, P., Conte, S., Compostela, C., & Guerani, L. (1993). Amodal completion in mouse vision. *Perception*, *22*, 713–721.

Kellman, P. J., & Spelke, E. S. (1983). Perception of partly occluded objects in infancy. *Cognitive Psychology*, *15*, 483–524.

Kitaoka, A., & Ishihara, M. (2000). Three elemental illusions determine the Zöllner illusion. *Perception & Psychophysics*, *62*, 569–575.

Kitayama, S., Duffy, S., Kawamura, T., & Larsen, J. T. (2003). Perceiving an object and its context in different cultures: A cultural look at New Look. *Psychological Science*, *14*, 201–206.

Lazareva, O. F., Levin, J. L., Vecera, S. P., & Wasserman, E. A. (2006). The search for object-based attention in pigeons: Failure and success. In K. Fujita & S. Itakura (Eds.), *Diversity of cognition: Evolution, development, domestication, and pathology* (pp. 3–37). Kyoto, Japan: Kyoto University Academic Press.

Lea, S. E. G., Slater, A. M., & Ryan, C. M. E. (1996). Perception of object unity in chicks: A comparison with the human infant. *Infant Behavior and Development*, *19*, 501–504.

Malott, R. W., Malott, M. K., & Pokrzywinski, J. (1967). The effects of outward-pointing arrowheads on the Müller-Lyer illusion in pigeons. *Psychonomic Science*, *9*, 55–56.

Morinaga, S. (1956). An examination of the conditions determining size-contrast. Paper presented at the 20th annual meeting of the Japanese Psychological Association, Tokyo, July.

Nagasaka, Y., & Osada, Y. (2000). Subjective contours, amodal completion, and transparency in animals. *Japanese Journal of Animal Psychology*, *50*, 61–73. (in Japanese with English summary)

Nakamura, N., Fujita, K. Ushitani, T., & Miyata, H. (2006). Perception of the standard and the reversed Müller-Lyer figures in pigeons (*Columba livia*) and humans (*Homo sapiens*). *Journal of Comparative Psychology*, *120*, 252–261.

Nakamura, N., Watanabe, S., Betsuyaku, T., & Fujita, K. (2010). Do bantams (*Gallusgallus domesticus*) experience amodal completion? An analysis of visual search performance. *Journal of Comparative Psychology*, *124*, 331–335.

Nakamura, N., Watanabe, S., Betsuyaku, T., & Fujita, K. (2011). Do bantams (*Gallus gallus domesticus*) amodally complete partly occluded lines? An analysis of line classification performance. *Journal of Comparative Psychology*, *125*(4), 411–419.

Nakamura, N., Watanabe, S., & Fujita, K (2008). Pigeons perceive the Ebbinghaus–Titchener circles as an assimilation illusion. *Journal of Experimental Psychology: Animal Behavior Processes*, *34*, 375–387.

Nakamura, N., Watanabe, S., & Fujita, K. (2009). Further analysis of perception of the standard Müller-Lyer figures in pigeons

(*Columba livia*) and humans (*Homo sapiens*): Effects of length of brackets. *Journal of Comparative Psychology*, *123*, 287–294.
Nakamura, N., Watanabe, S., & Fujita, K. (2014). A reversed Ebbinghaus-Titchener illusion in bantams (*Gallus gallus domesticus*). *Animal Cognition*, *17*, 471–481.
Obonai, T. (1954). Induction effects in estimates of extent. *Journal of Experimental Psychology*, *47*, 57–60.
Oyama, T. (1975). Determinants of the Zöllner illusion. *Psychological Research*, *37*, 261–280.
Parron, C., & Fagot, J. (2007). Comparison of grouping abilities in humans (*Homo sapiens*) and baboons (*Papio papio*) with the Ebbinghaus illusion. *Journal of Comparative Psychology*, *121*, 405–411.
Rauschenberger, R., & Yantis, S. (2001). Masking unveils preamodal completion representation in visual search. *Nature*, *410*, 369–372.
Regolin, L., & Vallortigara, G. (1995). Perception of partly occluded objects by young chicks. *Perception & Psychophysics*, *57*, 971–976.
Révész, G. (1924). Experiments on animal space perception. *British Journal of Psychology*, *14*, 386–414.
Robinson, J. O. (1972). *The psychology of visual illusion*. Mineola, NY: Dover.
Sato, A., Kanazawa, S., & Fujita, K. (1997). Perception of object unity in a chimpanzee (*Pan troglodytes*). *Japanese Psychological Research*, *39*, 191–199.
Sekuler, A. B., Lee, J. A. J., & Shettleworth, S. J. (1996). Pigeons do not complete partly occluded figures. *Perception*, *25*, 1109–1120.
Sovrano, V. A., & Bisazza, A. (2008). Recognition of partly occluded objects by fish. *Animal Cognition*, *11*, 161–166.
Sugita, Y. (1999). Grouping of image fragments in primary visual cortex. *Nature*, *401*, 269–272.
Suzuki, K. (1995). *Doobutsu wa sekai wo doo miruka* [*How animals see the world*]. Tokyo: Shinyo-sha. (in Japanese)
Timney, B., & Keil, K. (1996). Horses are sensitive to pictorial depth cues. *Perception*, *25*, 1121–1128.
Torii, S., & Mochizuki, T. (1995). Post-surgery perception of solids in the cases of the congenitally blind. *Japanese Psychological Review*, *38*, 581–605. (in Japanese with English summary)
Tvardíková, K., & Fuchs, R. (2010). Tits use amodal completion in predator recognition: A field experiment. *Animal Cognition*, *13*, 609–615
Ushitani, T., Fujita, K., & Yamanaka, R. (2001). Do pigeons (*Columba livia*) perceive object unity? *Animal Cognition*, *4*, 153–161.
Warden, C. J., & Baar, J. (1929). The Müller-Lyer illusion in the ring dove, *Turtur risorius*. *Journal of Comparative Psychology*, *9*, 275–292.
Watanabe, S. (2001). Discrimination of cartoons and photographs in pigeons: Effect of scrambling of elements. *Behavioural Processes*, *53*, 3–9.
Watanabe, S., Nakamura, N., & Fujita, K. (2011). Pigeons perceive a reversed Zöllner illusion. *Cognition*, *119*(1), 137–141.
Watanabe, S., Nakamura, N., & Fujita, K. (2013). Bantams (*Gallus gallus domesticus*) also perceive a reversed Zöllner illusion. *Animal Cognition*, *16*, 109–115.
Winslow, C. N. (1933). Visual illusions in the chick. *Archives of Psychology*, *153*, 1–83.
Yanagisawa, N. (1939). An experimental study on a modified Müller-Lyer figure (1). *Japanese Journal of Psychology*, *14*, 321–326.

Chapter 4
An Analysis of Theoretical Approaches to Geometrical-Optical Illusions

Barbara Gillam

DEFINITION OF ILLUSIONS

It is not altogether easy to define the term "illusion." It carries the connotation of being fooled into not seeing the real situation. But it does not seem correct to say that any visual experience not in accord with physical reality is an illusion. This would include much of Western painting, which creates a three-dimensional (3D) impression on a two-dimensional (2D) surface. We do not call such paintings illusory, however, because we do not feel fooled by them; we recognize how they achieve their effects by mimicking the perspective and occlusion information that normally arises from viewing a 3D scene. Illusions, on the other hand, are visual effects that we do not understand. For over 150 years the geometrical optical illusions, such as those in Figure I.4-1, have been considered prime examples: simple line drawings producing very large perceptual deviations from their measured geometric properties. These illusions are impervious to cognitive knowledge about the error. They appear to be based on automatic visual processes. Yet despite the long-standing preoccupation with the geometrical-optical illusions in vision science, and the innumerable papers devoted to them (with good summaries by Coren & Girgus, 1978, and Robinson, 1972/1998), they still have no agreed-upon explanations. They are the illusions par excellence. This chapter attempts to present an overview of their current theoretical status.

THE SPECIAL ROLE OF ILLUSIONS IN PERCEPTUAL THEORY

It is difficult to discriminate between theories of vision on the basis of how well they predict realistic perception, since all theories are designed to predict the same outcome. As deviations from reality, illusions hold the promise of supporting one particular theory and not others. They have been attributed to a wide range of visual processes. Some of these are ad hoc, specifically designed to explain a particular illusion, adding little to understanding. More promising are explanations in which illusions are explained as byproducts of processes considered to be useful in other contexts. Each new type of visual theory has tended to be eventually applied to illusions.

TWO VISUAL SYSTEMS AND TRANSITIVITY AMONG VISUAL JUDGMENTS

Geometric illusions provide a useful way of examining the relationships among different perceptual responses. Does the perceived deviation from reality that is an illusion apply to a motor response, for example pointing or eye movements, and does it entail deviations in the other visual properties that in a Euclidean world would be entailed? These issues are important theoretically in shedding light on how perception is organized. Initial comparisons of perceptual and motor responses found that the Ebbinghaus illusion of perceived size (Fig. I.4-1) did not seem to influence the motor response of reaching. This was taken as significant corroboration of a view, arising from studies of the effects of brain injury, that action and perception involve different streams in the visual cortex, with recognition based on the ventral stream while the dorsal stream is concerned with action (Goodale & Milner, 1992). A corollary of this view is that perception depends on responding to stimuli in world-centered (allocentric) coordinates while action requires body-centered (egocentric) coordinates. This distinction does not, however, neatly divide perception and action. Successful grasping depends on real size, thus allocentric coordinates, while eye movements, also a motor response, depend on oculocentric coordinates. Recent studies summarized by Bruno and Frantz (2009) mostly using the Müller-Lyer figure (Fig. I.4-1) show that contrary to earlier reports, grasping does exhibit an illusion similar to perceptual judgments when on-line feedback is precluded. Eye movements (saccades) to the Müller-Lyer figure also show an illusion significantly greater than chance but much less than shown in perceptual judgments. The magnitude of the effect depends on conditions such as exposure duration of the figure and starting position of the saccade. The Goodale and Milner theory has led to prolific research on the action/perception issue using geometrical-optical illusions producing very mixed results. This has led to the conclusion that different components of action or kinds of action may themselves have different stimulus requirements (Smeets, Brenner, de Grave, & Cuijpers, 2002). It has also been found that there are intransitivities or inconsistencies among perceptual judgments, such as size and position, so that when a motor and a perceptual response to a figure differ, it may not

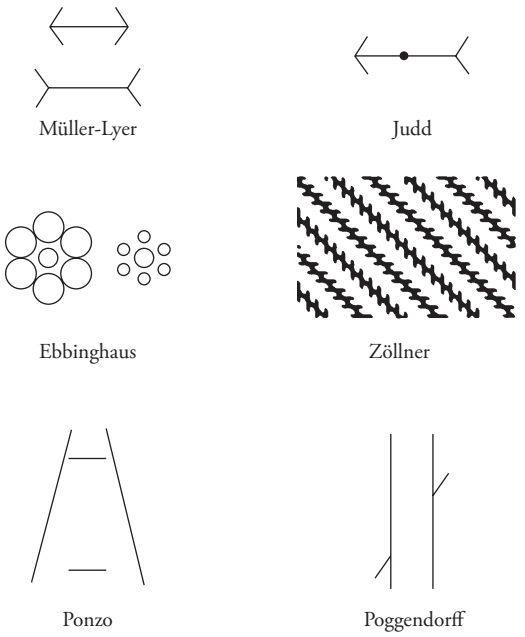


Figure I.4-1. Some of the most famous geometrical-optical illusions. In the Müller-Lyer the horizontal lines are the same length. In the Judd the dot is centered on the horizontal line. In the Ebbinghaus the central circles are the same size. In the Zöllner the oblique lines are parallel. In the Ponzo the horizontal lines are equal length. In the Poggendorff the oblique lines are collinear. From a variety of earlier papers by the author.

be because the action is carried by a different visual stream but because it is based on a different perceptual property. For example, the Müller-Lyer illusion, which is a size illusion, does not entail a corresponding position illusion when observers are asked to point to the vertex with an unseen hand (Mack, Heuer, Villardi, & Chambers, 1985). However, observers also fail to make significant position errors when asked to set a computer probe vertically beneath the vertices of the Müller-Lyer figure (Gillam & Chambers, 1985), even though this is a perception task and in a Euclidean world size and position would map onto one another. Each perceptual judgment, like each motor response, seems to be based on specific information relevant to it. Illusion findings tell us that we should not think of perception as being read off locations on a Euclidean (or any other) map or representation of a scene (Gillam 1998; MacLeod & Willin, 1995). This makes explaining illusions extremely difficult because we may be in error if we try and explain them via distortions in a related Euclidean property.

It has also been claimed that geometrical-optical illusions are special in being absent under conditions of isoluminance in which a figure differs from its background only in color. Recent measurements, however, have shown that illusions are present under these conditions and thus not restricted to the luminance processing system (Hamburger, Hansen, & Gegenfurter, 2007).

MAJOR VARIETIES OF ILLUSION THEORY

Contrast and Assimilation

These are modern versions of 19th-century theories that do not go far beyond describing the illusion. Assimilation refers to the unrecognized influence of the perception of whole figures on the perception of their parts—as when the shaft length of the Müller-Lyer illusion (see Figure I.4-1) is said to assimilate to the length of shaft plus attachments (although this idea seems to apply better to the outgoing fin version than the ingoing fin version). Contrast refers to a situation in which part of a figure appears to have an exaggerated difference from another part of the figure. The Zöllner illusion (Fig. I.4-1) is an example. The long parallel lines appear to diverge with each one deviating in orientation from the orientations of the small oblique lines that cut across them. Insofar as assimilation and contrast theories go beyond simply describing the illusion, they belong either in the category of physiological theories or functional theories and are discussed further under these categories.

Physiological Theories

Illusions are often attributed to physiological processes. Sometimes these are well-informed theories, but sometimes competent visual physiologists make forays into explaining an illusion via this or that physiological process with little knowledge of the detail and complexity of illusion findings over the past 150 years.

The Müller Lyer illusion has been attributed (Ginsberg, 1979) to the influence of cortical low-pass filtering on size judgments (a reliance on coarse spatial frequency channels). Fins and shaft would be blurred in those channels. This explanation is not well supported, however. The illusion persists when the lines are composed of balanced dots, which carry only high spatial frequency information (Carlson, Moeller, & Anderson, 1984) although Morgan (1995) has argued there could be second-stage filtering causing blur in cortical low spatial frequency channels.

Contrast, as in the Zöllner illusion, has been attributed to retinal blur at acute angles or more recently to cortical cross-orientation inhibition between lines of similar orientation (Blakemore, Carpenter, & Georgeson, 1970). Purely retinal processes have been ruled out as the basis of this and most other geometrical-optical illusions by the finding that they occur in "Julesz random dot stereograms," which are images composed of random dots in the two eyes from which forms emerge only in the fused image that is created cortically. There is, however, evidence (Glennerster & Rogers, 1993) that at least part of the distortion in the Müller-Lyer illusion may occur before fusion, because if the outgoing fin figure is presented to one eye leading to a width enhancement in that eye's view, stereoscopic slant may be seen based on binocular disparity with an unenhanced image in the other eye.

Cross-orientation inhibition is based on the idea that since any given orientation is coded by the population responses of cells sensitive to a range of orientations, a bias could be introduced by the inhibitory effect on the population response by the presence of a closely neighboring orientation. This kind of theory applied to the Zöllner illusion has difficulties with the finding that the parallel lines still show a tilt offset when replaced by aligned spaces (Fig. I.4-2) or when the short "inducing" lines have opposite contrast polarity from the parallel lines (Earle & Maskell, 1995). In the former case, however, there is a suggestion of subjective contours along the edges of the spaces, and it is known in other illusion figures such as the Poggendorff

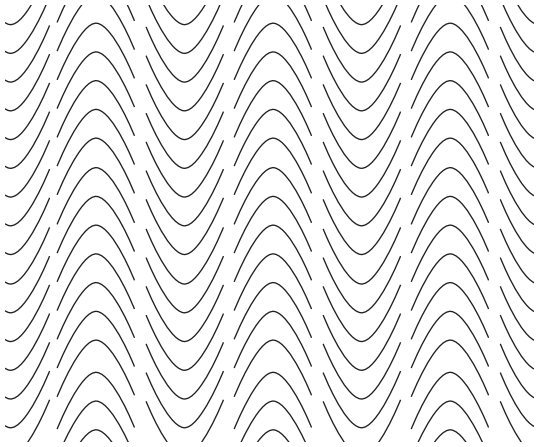


Figure I.4-2. A version of the Zöllner illusion in which the oblique lines are replaced by spaces. The illusion persists, showing that physical angles are not critical. Earle and Maskell (1995) (with permission from *Perception*).

(Fig. I.4-1) that an illusion is still present when luminance contours are replaced by subjective ones. Whether the major physiological models can be applied to illusions formed by subjective contours depends on what one thinks subjective contours are.

The Poggendorff illusion (Fig. I.4-1) refers to an apparent noncollinearity in the collinear transverse oblique lines. This has also been attributed to inhibitory interactions between the representations of the transversals and parallels, which cause the transversal orientation to be shifted away from the orientation of the parallels. Unlike the Zöllner illusion, however, the Poggendorff illusion is destroyed if the transversals cut across the parallels, and it is optimum for angles much greater than those at which the Zöllner occurs. Also, the illusion is destroyed when only the acute angle components survive (Fig. I.4-3a). These should demonstrate at least as much cross-orientation inhibition as the obtuse angle components, which have an illusion almost as strong as the full effect (Fig. I.4-3b). Theories based on angle enhancement are also inconsistent with the finding that the apparent orientation of the transversals in a Poggendorff configuration does not differ from the orientation of the same lines in isolation and that the illusion is reduced almost to zero if the entire figure is rotated so that the transversals are horizontal or vertical. Morgan (1999)

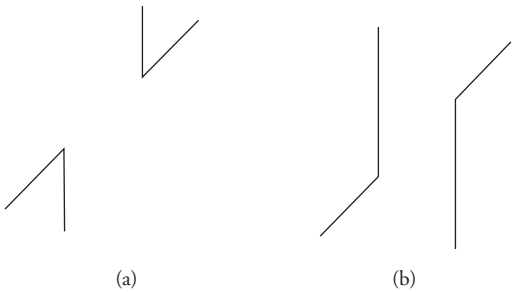


Figure I.4-3. Amputated versions of the Poggendorff illusion. When only the acute angle parts are retained (a), the illusion disappears. When only the acute angle components are retained (b), the illusion persists with obtuse angles. From Gillam (1998).

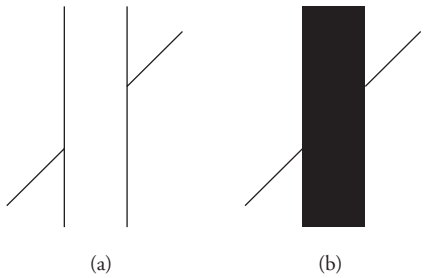


Figure I.4-4. The Poggendorff illusion in (a) is greatly diminished when the transversals and the area between them share the same contrast with the background (b). The transversals in (b) appear to modally complete in front of the central rectangle with a much diminished displacement. (Modified from figures in Spehar & Gillam, 2002.)

attributes the Poggendorff illusion to blur in second-stage filters resulting in a shifting of the intersections of the transversals and parallels into the acute angles. He argues that this changes the orientation of the implicit line between the transversals while not changing the orientation of the transversal themselves: hence the apparent noncollinearity. There is evidence for a difference in perceived orientation between the implicit and explicit oblique lines in the Poggendorff figure (Day & Dickinson, 1976), but this cannot result from a displacement of the position of the intersections. There is no illusory displacement when observers are asked to set a moveable dot next to the Poggendorff figure so that it is horizontally aligned with the intersection. The Morgan theory also has problems accounting for the huge effects of figure rotation and the amputations shown in Figure I.4-3, unless it is argued that second-stage blur only influences vernier-type judgments for complete Poggendorff figures at certain orientations.

Another observation strongly at odds with physiological theories of the Poggendorff illusion can be seen when the parallels are replaced by a solid rectangle of the same luminance and color as the transversals (Fig. I.4-4b). This should not change the predictions of theories based on angle enhancement or blur. Yet the illusion is markedly diminished in this case relative to its regular outline form (Fig. I.4-4a). The transversals appear to complete across the front of the solid rectangle as if camouflaged against it. This is called "modal completion" and is not possible for the outline figure. Spehar and Gillam (2002) showed that the strength of the illusion was reinstated, however, when the solid rectangle and the transversals were of opposite contrast polarity relative to the background (e.g., with the rectangle black and the transversals white against a grey background), making modal completion impossible. These observations are more consistent with illusion theories based on spatial layout (see later discussion) than with physiological theories.

Apart from the problems physiological theories have in explaining particular illusions, they are subject to a conceptual difficulty. Sometimes these theories seem to imply that it is the image that we see rather than the environment, so that all that is required in a theory of an illusion is a mechanism-producing image distortion (retinal or cortical). However, surely perception would calibrate whatever representation exists to provide an accurate perception of

the external world, which is its task. In fact a number of studies, mostly but not exclusively using the Müller-Lyer figure, have shown that illusions diminish strongly with massed exposure (Coren & Girgus, 1978). This suggests that they are reinforced in everyday viewing. To explain illusions, it does seem necessary to show what function they would have in normal vision.

Functional Theories

Functional theories have always been popular because illusions are of such magnitude that it seems likely that they serve a function in normal vision. Functional theories can, of course, coexist with physiological theories, although the emphasis is usually on one or the other. Morgan (1995) has argued that illusions arise when the visual system is required to perform a task to which it is ill adapted. On this basis Morgan, Hole, and Glennerster (1990) proposed a functional theory of size assimilation illusions such as the Müller-Lyer. They argue that the visual system is well adapted to making metrical judgments based on the average positions of clusters of features but ill adapted to making metric judgments between arbitrarily abstracted features within clusters. They argue that although we need to encode the distance between objects to navigate through the environment, it would be too cumbersome to encode the distance between every pair of features in the image. There may instead be a hierarchical system by which object relations are encoded and then relations within objects if required. This argument is supported by their finding that judgments of the distance between two dot clusters is based on their centroids and that the intercentroid distance tends to bias judgments between a designated dot in one cluster with one in another cluster. This theory (which, like most illusion theories, has precursors in the 19th century) can accommodate the fact that the Müller-Lyer size difference occurs with appendages other than fins, for example circles and boxes (Fig. I.4-5). However, it does not easily accommodate the fact that even a very small displacement between the fins and the shafts markedly reduces the Müller-Lyer illusion (Worrall & Firth, 1974). It also does not explain why the fin form of the illusion is considerably stronger than forms with other appendages. Also, in the experiment comparing dot separations across clusters, these authors found the least bias for dots at the edges of the clusters. The outgoing fin form of the Müller-Lyer resembles this case in that the ends of the shaft coincide with the endpoints of the fins. Yet the outgoing fin figure produces a very strong effect relative to a plain line—stronger than the ingoing fin form. Morgan (1995) makes the valid point, however, that the Müller-Lyer illusion will not be understood until size perception is understood more generally. I would tend to add that to understand illusions it is necessary to understand how we decode projections of the 3D world, which brings us to the most famous and persistent of the functional theories.

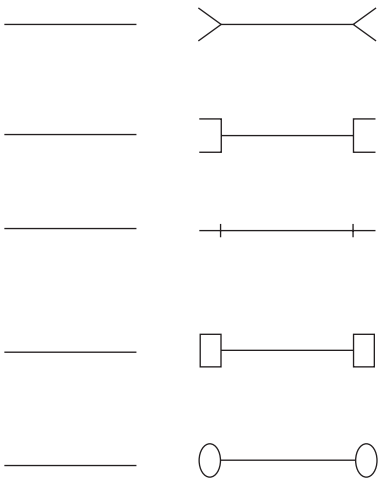


Figure I.4-5. Versions of the Müller-Lyer illusion with different appendages. In all versions the lengths of the central horizontal lines and the overall figure widths are the same. The degree of illusion varies. (From Gillam, 1998.)

It is generally acknowledged that the visual system is evolved/adapted to respond to 3D scenes not 2D line drawings. Theories dating back to the 19th century (when most of the geometric illusions were discovered) point out that many illusions can be thought of as serving 3D scene perception despite not appearing to be 3D. The claim is that the "illusions" that occur in sketchy line drawings would often represent veridical perception if the same deviations from image properties were seen in the 3D world. Figure I.4-6 shows a perspective drawing of a room in which three well-known illusions are embedded: the Müller Lyer, the Poggendorff, and the Ponzo. If this were an actual 3D scene, an observer would not know what one was talking about if one asked if the picture rail is collinear with the baseboard (although they are collinear in the image). Perception is not geared to detecting properties of the image, and it is very difficult to make a judgment of projective (2D) properties such as relative retinal size and collinearity when looking at a 3D scene. However the *picture* of a scene (as in Figure I.4-6) has a dual phenomenological existence. It can readily be seen both as a 2D arrangement and as a 3D scene. Viewing it as a 2D arrangement, it seems obvious that the baseboard is not projectively collinear with the picture rail. This erroneous judgment is clearly influenced by crosstalk from the 3D interpretation of the picture, in which the baseboard appears to be on a different horizontal plane from the picture rail. Finally, if more context is removed, the same configuration of projectively collinear lines attached to the opposite sides of a rectilinear surface no longer appears 3D at all. The judgment appears to be based on a purely 2D representation, and the apparent noncollinearity is called the Poggendorff illusion. But is it unrelated to the 3D processing we saw in the real scene and the picture? Or is crosstalk from 3D processing, which we are now unaware of, still occurring? That is the claim of 3D processing theories of illusions. The most famous of these is Gregory's (1970) "constancy scaling" theory. Gregory's photographs of near and far corners mimicking the ingoing and outgoing fins of the Müller-Lyer illusion is unfortunate since it has been widely taken to imply that one kind of corner is seen as larger because it is seen as further. Many critics have pointed out that the relative distances of corners is arbitrary and that similar illusions occur within a single figure, such as the Judd illusion (Fig. I.4-1) in which the midpoint of the shaft is displaced away from

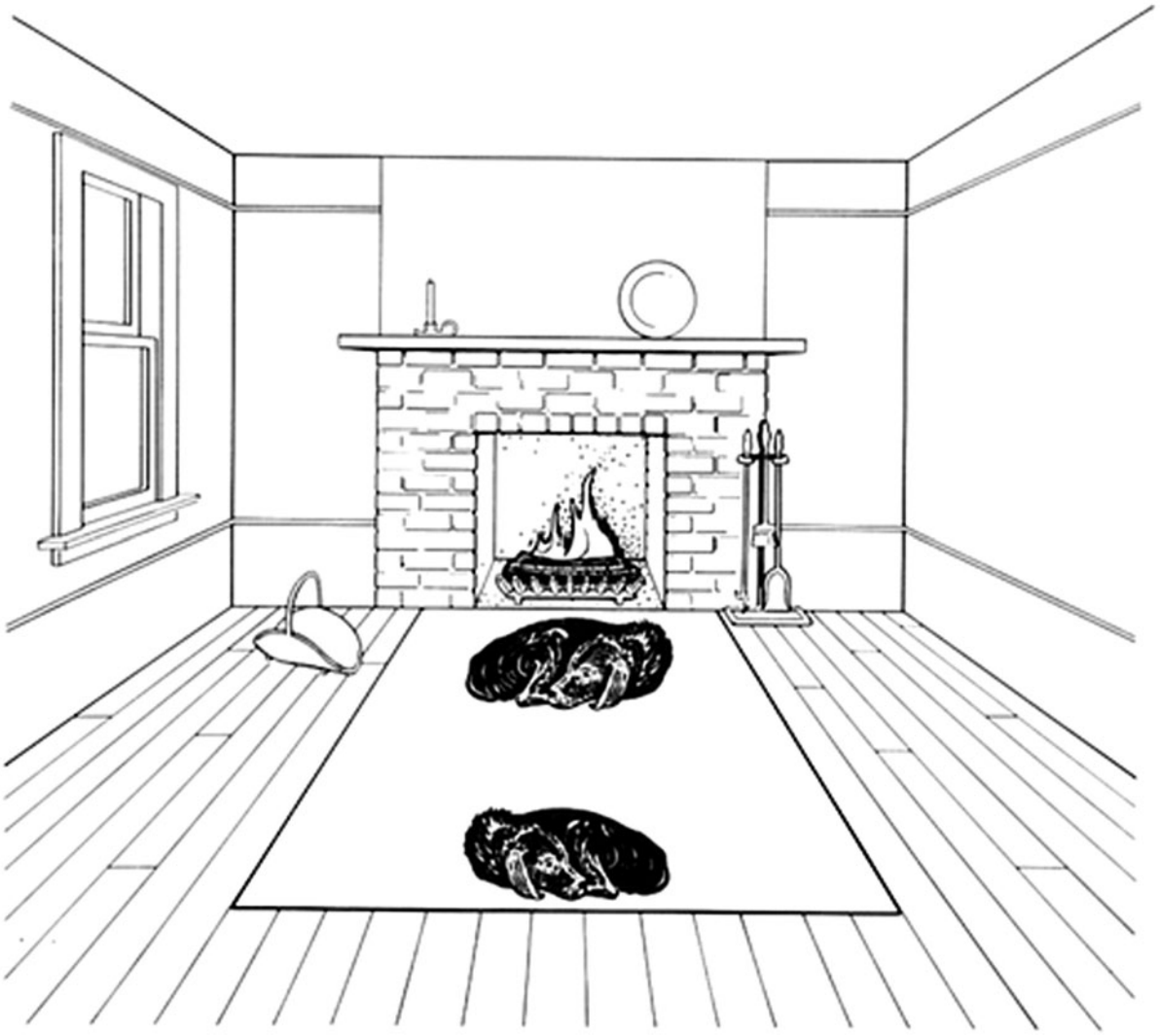


Figure I.4-6. Three of the familiar illusions embedded in a scene. The front of the carpet is the same length as the bottom of the back wall (partial Müller-Lyer illusion). The base board on the left is collinear with the picture rail on the right (Poggendorff illusion). The dogs are the same size (Ponzo illusion). (From Gillam, 1998.)

the obtuse angle and toward the acute angle. The idea of unconscious depth, assumed to be necessary to Gregory's theory, has also been much criticized. To do Gregory justice, he did propose a form of size scaling he called "primary size scaling." This term refers to scaling that is the direct result of stimulus context and is not mediated by apparent depth. However, Gregory did not develop this idea. His photographs of the two corners, his experiments manipulating apparent depth, and some comments in his writings have resulted in his many critics disregarding the concept of primary constancy scaling or treating it as a post hoc theory applying only to illusions. However, this is not a fair criticism. Even in regular (nonillusory) perception, the idea that perceived size is based on a calculation using retinal size and perceived distance (the size-distance invariance hypothesis) has been refuted many times. (For a detailed account of this literature see Sedgwick, 1986). The view that perceived size is directly given by stimulus information and not calculated from perceived depth goes back at least as far as Bishop Berkeley (1709), who pointed out that perceived size and perceived distance are based on similar "ideas" (by which he seems to mean what would now be called "cues"). He states

> They do not first suggest distance, and then leave it to the judgment to use that as a medium whereby to collect the magnitude; but they have as close and immediate a connection with the magnitude as with the distance; and suggest magnitude as independently of distance as they do distance independently of magnitude. (para. 53)

Gibson (1966) has made a similar point with his direct theory of perception based on the discrimination of relationships ("invariants") in the "optic array" (the pattern of light arriving at a particular vantage point). Size and distance may use similar information in different ways, for example weighing it differently or requiring different supporting information.

For a good example of how size scaling can work directly on the basis of contextual information and not via perceived distance, consider Figure I.4-7a, the conventional Ponzo

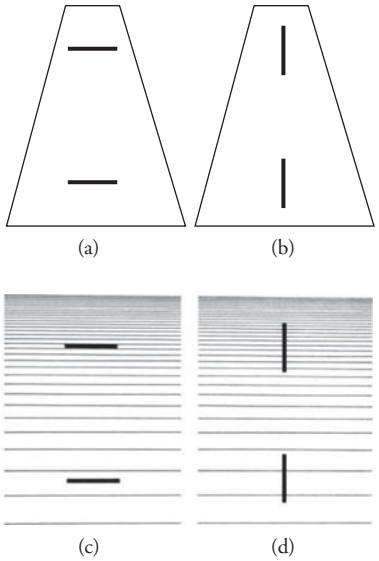


Figure I.4-7. (a) is the classical Ponzo illusion. (b) shows the same figure with the horizontal lines rotated to the vertical. The illusion disappears. (c) shows the horizontal lines placed on a ground-plane perspective consistent with receding depth, but there is no size illusion. In (d) the horizontal lines in (c) are rotated to the vertical. Now a Ponzo-type illusion is seen. The illusion does not depend on perceived depth but on the background size scale. (Based on figures in Gillam, 1998.)

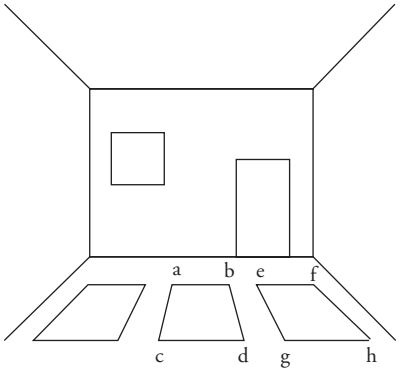


Figure I.4-8. A diagram of scene perspective in which acute and obtuse angles have a variety of near and far locations. Acute angles are not always near and obtuse angles far. (From Gillam, 1998.)

illusion in which the upper test line appears wider than the lower test line. Figure I.4-7b shows the same context but with the test lines rotated 90°. The illusion now disappears, showing that it cannot be due to the perceived depth (conscious or unconscious) produced by the linear perspective of the converging lines. However, Figure I.4-7c and I.4-7d reverse this situation. Now it is the vertical lines that have the illusory size difference, while there is no illusion for the horizontal lines. Why is this the case? The reason is that the background scale in Figure I.4-7c and I.4-7d is of vertical extents, and the illusion is seen in the vertical test lines. In Figure I.4-7a and I.4-7b, the background scale is of horizontal extents, hence the illusion affects only lines in that orientation. This is a good example of primary constancy scaling, with size based on the relationship of test objects to an appropriate background size scale. Other size contrast illusions such as the Ebbinghaus illusion (Fig. I.4-1) have a similar explanation. It seems obvious that this kind of size scaling must be ubiquitous in normal viewing. Scaling size relative to an appropriate background scale automatically compensates for distance while not depending on perceived distance.

Can the Müller Lyer illusion (Fig. I.4-1) be explained by contextual size information? The most important fact about the Müller-Lyer is that it is actually quite a local effect. The apparent size of a line is enhanced by a single attached obtuse angle and diminished by a single attached acute angle. The local nature of the effect of angles can be seen in the Judd figure (Fig. I.4-1). Also, the size illusion has been shown to occur only for the portion of the shaft that is close to the angle. Indeed, the size of the illusion has been found to increase with the number of the angular attachments to the shaft (Tausch, 1954). The full illusion has four acute angles attached to one line shaft and four obtuse angles attached to the other and can be regarded as a super stimulus for size scaling, or what Morgan and Casco (1990) have called all illusions, namely effects evolved in the literature not to be informative but to be conspicuous.

Is there any aspect of 3D viewing that is consistent with an influence of angles on local line length? Figure I.4-8 shows a picture of 3D rectilinear surfaces whose edges and corners incorporate the Müller-Lyer and Judd configurations. It is clear that both acute and obtuse angles in the image can arise from either near or far corners on rectilinear surfaces, depending on the lateral position of the surface in the field of view. However, regardless of whether near or far, acute angles obtrude from their projected shapes while obtuse angles intrude into the shapes. Seeing the parts of the lines bounded by acute angles as shortened and those bounded by obtuse angles as lengthened would tend to make the surfaces more rectilinear, as they usually are in the 3D "carpentered" environment. It has been suggested (Warren & Bashford, 1977) that the angle-based scaling of the Müller-Lyer illusion is associated with shape constancy rather than size constancy. This is a tricky argument because remember there is no local position illusion associated with the Müller-Lyer illusion (although to complicate matters two opposite Judd figures appear displaced as a whole). For a discussion of such paradoxes, see Gillam (1998).

The fact that the illusion occurs for appendages other than fins. (Fig. I.4-5) has been used as a prime argument against 3D processing theories of the Müller-Lyer illusion, which are based on angles. However, the size illusion produced by replacing the fins of the Müller-Lyer figure with open boxes is only about a quarter as great as the one with fins (Tausch, 1954). Closed boxes have up to half the effect of the Müller-Lyer, but they introduce another factor not present in the classical Müller Lyer. They appear to occlude the shaft, and occlusion has a known lengthening effect on the apparently occluded object (Kanizsa, 1979).

The Poggendorff illusion is embedded in a number of 3D scene configurations in which the transversals represent horizontal receding lines (Figs. I.4-6 and I.4-9). Gillam (1971, 1998) and Green and Hoyle (1963) proposed that the illusion is due to crosstalk from such 3D interpretations. Gillam's theory is illustrated in Figure I.4-10. It assumes that oblique lines in a 3D context are processed as receding lines and that two spatially separated but aligned lines of

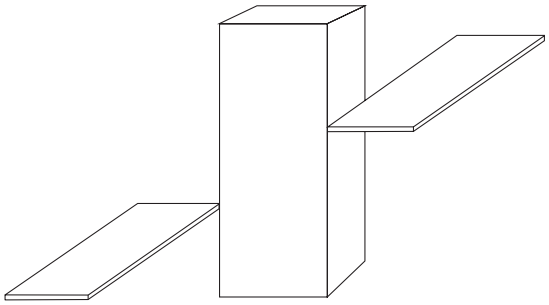


Figure I.4-9. The Poggendorff illusion in a realistic context showing that collinearity in the image is not an indication of collinearity in the scene depicted. (Based on a figure from Rock, 1975.)

the same oblique orientation are processed as collinear in a single receding plane of 3D space with the height difference between their laterally displaced inner ends processed as a depth difference. This situation is illustrated on the left side of Figure I.4-10 and is consistent with collinearity of *ab* and *cd* in 3D space. However, when the inner ends of the lines (*bc*) are attached to a frontal plane, as in the Poggendorff figure, their height difference is no longer consistent with a depth difference between the lines but with a physical difference in height where they are attached to the frontal plane. This arrangement is illustrated on the right side of Figure I.4-10. It is not consistent with collinearity of the lines *ab* and *cd* in 3D space. This theory of the basis of the Poggendorff illusion was tested (Gillam, 1971) by changing the projected plane between the transversals to a receding plane with the same vanishing point as the transversals keeping conditions on which the illusion normally depends, such as the orientation of the transversals and the width between the parallels, unchanged. All eight observers showed a marked reduction in degree of illusion under these conditions with a reduction of the mean illusion from 13 to 6 mm. Figure I.4-11 demonstrates this effect for multiple collinear oblique lines with a common vanishing point. Gillam's finding strongly supports the theory that the consistency of the image with collinearity of the transversals in 3D space is an important factor in the Poggendorff illusion. Others have also supported this idea experimentally (e.g., Konig & van Lier, 2007). The 3D theory accounts for the diminution of the illusion

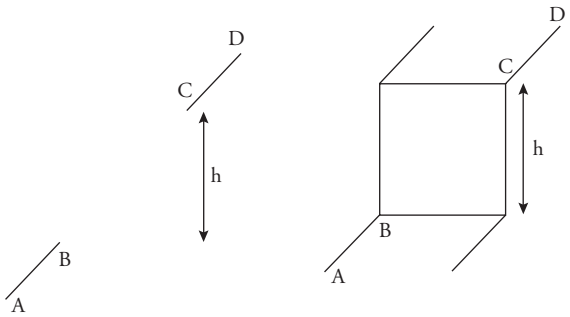


Figure I.4-10. Illustrating Gillam's theory of the Poggendorff illusion. The same oblique lines (*ab* and *cd*) can be either collinear in 3D space (left figure) or on different planes as indicated by their attachment to a frontal plane (right figure). (From Gillam, 1998.)

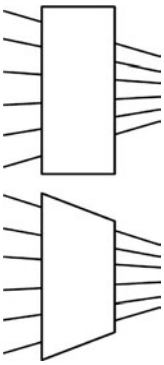


Figure I.4-11. The apparent collinearity of transversals attached to a central plane depends on their relationship of their vanishing point to the vanishing point of the plane. In B the transversals and central plane have the same vanishing point and are consistent with collinearity in 3D space. Poggendorff displacements are much diminished relative to A, where the central rectangle does not share a vanishing point with the transversals. (From Gillam, 1998.)

with the acute angle components alone and its persistence with the obtuse angle components alone by treating the parallel components in these two cases as planes (Gillam, 1971). The near elimination of the illusion when the entire Poggendorff figure is rotated so that the transversals are vertical or horizontal is consistent with the theory, since the transversals in this case would no longer be processed in 3D space as receding lines.

As predicted, the Poggendorff illusion is strongly reduced if the transversals are separated in stereoscopic depth from the parallels, especially if they are in front. Also consistent with a scene-based theory is a related finding, mentioned earlier, that modal completion of the transversals across the parallels greatly reduces the illusion by allowing the transversals to be continuous in the scene.

It should be mentioned, however, that there are several findings that are not easy for the 3D processing theory of the Poggendorff illusion to account for. It can occur sometimes in the real world, for example when a line is strung obliquely and not in depth behind a column. The illusion is clearly an automatic response to the configuration, not its cognitive interpretation. Figure I.4-12 (attributed to Horrell by Morgan, 1999) shows another problematic figure for 3D processing theory. The transversals are attached to what looks like a frontal plane, which they intersect at right angles. The illusion virtually disappears. This indicates an importance for local factors, which the theory does not encompass. On the other hand, Gillam (1971) showed that collinear lines attached at right angles to a trapezoidal figure can appear displaced if they are differently located

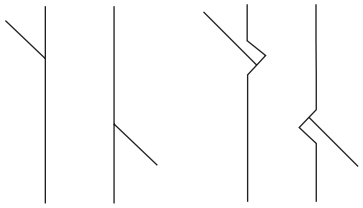


Figure I.4-12. Placing the transversals of a Poggendorff-type configuration so that they form local right angles eliminates the illusion. (From Morgan,1999—attributed to Horrell.)

relative to the vanishing point of the converging lines of the figure to which they are attached. It is the vanishing point that determines what is horizontal in 3D space.

Scene Statistics

Theories that explain illusions by crosstalk from 3D scene processing of the image would gain extra validity from scene statistics in which a correlation is established between properties of the image and 3D properties of the scene, which can now be obtained from laser scanning. For example, in the case of the Poggendorff illusion, it would be useful to determine if rectilinear planes intervening between the inner ends of collinear oblique lines in the image increase the probability that the latter are produced by noncollinear lines in the scene. This would be an example of what Geisler (2008) has called "cross domain" scene statistics (e.g., relations between image statistics and environment statistics). A good example of cross-domain statistics, although not related to geometric illusions, is the finding by Potetz and Lee (2011) of a correlation between pixel luminance and the range distance of that pixel in the scene, thus validating a claim by Leonardo da Vinci that luminance is a cue to distance. Cross-domain scene statistics are not easy to obtain and do not seem to have been carried out in the case of illusions. A systematic study might reveal many nonobvious cases in which projections resembling illusion figures derive from 3D arrangements. For example, Figure I.4-13a shows a Müller-Lyer -type configuration devised by DeLucia and Hochberg (1991) in which the space between acute angles appears less than the same space between obtuse angles. The authors use this figure to argue against an explanation of the Müller-Lyer illusion based on cross talk from 3D processing in that there is no obvious 3D origin for the size differences in their configuration. However, Figure I.4-13b (Gillam, 1998) shows that the DeLucia and Hochberg configuration could be the projection of a regular 3D arrangement of blocks with greater 3D linear extents associated with projective extents terminated at obtuse angles (*ca*) than for the same projective extents terminated at acute angles (*ab*), corresponding to the Müller-Lyer illusion.

The only major attempt to explain illusions by scene statistics obtained empirically has been undertaken by Howe and Purves (2005) as part of their "empirical theory of vision." Although they use range finders to validate that lines or planes in the image represent valid linear progressions in the scene, the statistics that form the basis of their theory are "within domain" (using Geisler's terminology)—essentially finding the probability distribution of one property in the image (e.g., length), given another image property (e.g., orientation). On this basis, Howe and Purves showed, for example, that the distribution of line lengths in scene images varies with line orientation. They argued that the perceived length of a particular line would depend on its percentile rank in the cumulative distribution of lengths for that orientation, not on its true length. Since scenes give rise to fewer long vertical line projections than long horizontal line projections, for example, a given vertical line will appear longer than a horizontal line of the same length having a higher percentile rank. So the bias is attributed to relations among image properties, which may or may not be related to the depths of their origins.

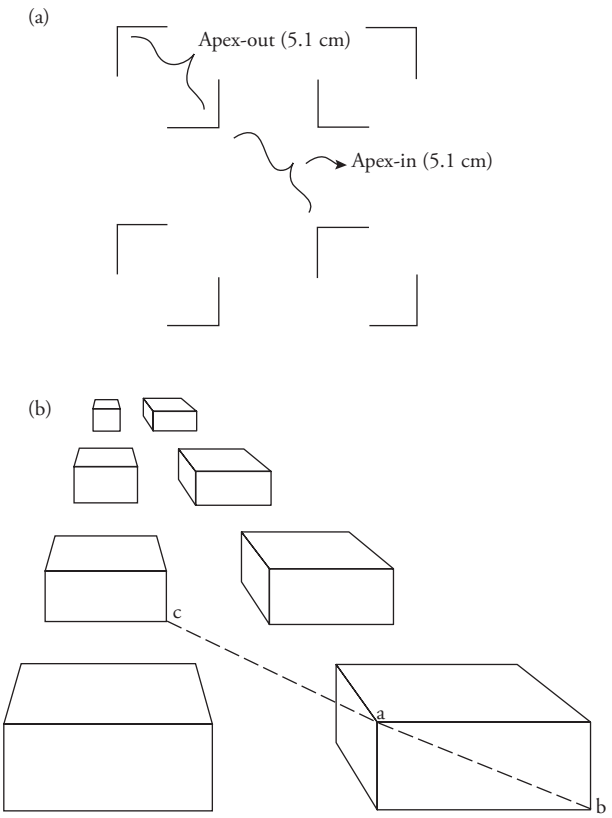


Figure I.4-13. (a) Drawing by DeLucia and Hochberg (1991), described in the text, showing how the Müller-Lyer illusion can be embedded in an apparently two-dimensional pattern with no obvious 3D interpretation. (Reproduced with permission of the Psychonomic Society.) (b) A figure showing how the same Müller-Lyer configuration as in A could arise from a 3D scene with the relevant image extents (*ca* and *ab*) representing different physical extents. (From Gillam, 1998.)

How is the Müller-Lyer illusion explained on this theory? First, locations in the scene were found where part of the Müller-Lyer configuration (either an ingoing or outgoing arrowhead on the left or right) could occur. Then neighboring regions were sampled consistent with valid locations of the shaft and complementary arrowhead for varying shaft lengths. (These need not be physically present lines.) The probability of occurrence of physical sources of Müller-Lyer projections as a function of the projected length of the shaft (or the interval between the two attachments) was tallied. They then determined how the probability distributions differed depending on the attachments. For an initial outgoing arrowhead on the left they found that the summed probability distribution of shaft and complementary arrowhead on the right was further to the right than for an initial ingoing arrowhead on the left. They argue on this basis that complementary outgoing arrowheads will appear displaced left relative to ingoing ones because they will be lower in the percentile rank of lengths for its attachment. Thus ingoing arrowheads will shorten the shaft and outgoing arrowheads will lengthen the shaft. The authors speculate that the reason for this is that since the sources of the

projections are straight lines, the configurations on which they lie are likely to be planes. For the outgoing arrows, the probability of the appendage falling off the plane occur at a greater separation than for the ingoing arrowheads. This will determine what inputs contribute to the probability weighting.

This theory bases the Müller-Lyer illusion on the probability of the joint configuration of complementary appendages at different separations. It has the advantage of finding similar biases for box appendages, but it fails to account for the fact that a single angle elicits one-quarter of the effect of the four angles and for the greater size effect of the obtuse angle than the acute angle figure. It also bases the illusion on biases in position, and the position of arrowhead vertices has been shown not to account for the size effect. The speculation that the probabilities obtained are based on the lines of the illusion lying on a plane is at odds with the frequency with which acute and obtuse angles in images occur at the intersection of different planes in the scene.

In their book, Howe and Purves (2005) argue that the Poggendorff illusion (Fig. I.4-1) and its variants reflect the "fundamental uncertainty of how the physical sources of projected line segments are actually continued across a spatial interval" (p. 87). The percepts elicited by these stimuli, they argue, should reflect the past experience of human observers with straight-line segments that are the same projected orientation but separated by a spatial interval. These could vary in position. Valid locations in scene images of the one transversal (say the left one) and the parallels are found. For each of these partial illusion configurations they sampled positions of projected lines attached to the right transversal of the same orientation as the left one but at different vertical locations. It was found that the probability distribution of such lines was above the continuation of a transversal pointing downward (toward the parallels) and above the continuation of one pointing up. According to the theory, this would result in the actual continuation of a transversal pointing downward toward the parallel lines appearing too low and the continuation of an upward pointing one appearing too high because of their percentile ranks in the distributions of lines in different positions but of the same orientation. One could question the assumptions underlying the choice of these scene statistics. It is not clear why the uncertainty of the continuation of an oblique line across a spatial interval should result in uncertainty about the vertical position of its emergence but not the orientation. This seems arbitrary. It is only because the illusion is one of position that this property is sampled while orientation is kept constant. As in the case of the Müller-Lyer illusion, the Poggendorff effect is explained by proposing that the existence of straight lines implies a planar surface in that region of the scene. Lines of the same orientation equally positioned above and below the true intersection point are closer to the center of the plane for downward-pointing arrows. Again, there is no acknowledgement that lines intersecting at an angle may not be on a plane. It is a weakness of this theory that it does not produce plausible bases for the statistical findings.

CONCLUSIONS

Illusions are very difficult to explain. Local physiological effects first come to mind but tend not to be supported when their predictions are examined in detail. Functional theories have produced promising data but also encounter observations that are difficult to account for. Scene statistics seem like a promising technique to pursue further, but it is not clear what one should be looking for.

REFERENCES

Berkeley, G. (1910). *A new theory of vision and other writings.* Everyman's Library London. J.M.Dent & Sons Ltd. (Original work published as *An Essay towards a new theory of vision*. (1709).

Blakemore, C., Carpenter, R. H. S., & Georgeson, M. A. (1970). Lateral inhibition between orientation detectors in the human visual system. *Perception, 9*(4), 467–474.

Bruno, N., & Franz, V. H. (2009). When is grasping affected by the Müller-Lyer illusion? A quantitative review. *Neuropsychologia*, *47*, 1421–1433.

Carlson, C. R., Moeller, J. R., & Anderson, C. H. (1984). Visual illusions without low spatial frequencies. *Vision Research, 24*(10), 1407–1413.

Coren, S., & Girgus, J. S. (1978). *Seeing is deceiving: The psychology of visual illusions*. Hillsdale, NJ: Lawrence Erlbaum.

Day, R. H., & Dickinson, R. G. (1976). The components of the Poggendorff illusion. *British Journal of Psychology*, *67*(4), 537–552.

DeLucia, P. R., & Hochberg, J. (1991). Geometrical illusions in solid objects under ordinary viewing conditions. *Perception & Psychophysics, 50*(6), 547–554.

Earle, D. C., & Maskell, S. J. (1995). Spatial filtering and the Zöllner-Judd geometrical illusion: Further studies. *Perception, 24*(12), 1397–1406.

Geisler, W. S. (2008). Visual perception & the statistical properties of natural scenes. *Annual Review of Psychology, 59*, 167–192.

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton-Mifflin.

Gillam, B. (1971). A depth processing theory of the Poggendorff illusion. *Perception & Psychophysics*, *10*(4A), 211–216.

Gillam, B. (1998). Illusions at century's end. In J. Hochberg (Ed.), *Perception and cognition at century's end* (pp. 95–136). New York: Academic Press.

Gillam, B., & Chambers, D. (1985). Size and position are incongruous: Measurements on the Müller-Lyer figure. *Perception & Psychophysics*, *37, 5*49–556.

Ginsburg, A. P. (1979). Visual perception based on spatial filtering constrained by biological data. *Proceedings of the International Conference on Cybernetics and Society.* New York: Institute of Electrical and Electronics Engineers.

Glennerster, A., & Rogers, B. J. (1993). New depth to the Müller-Lyer illusion. *Perception, 22*, 691–704.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience*, *15*(1), 20–25.

Green, R. T., & Hoyle, E. M. (1963). The Poggendorff illusion as a constancy phenomenon. *Nature, 200*, 611–612.

Gregory, R. L. (1970). *The intelligent eye.* London: Weidenfeld & Nicolson.

Hamburger, K., Hansen, T., & Gegenfurtner, K. R. (2007). Geometric-optical illusions at isoluminance. *Vision Research*, *47*, 3276–3285.

Howe, C. Q., & Purves, D. (2005). *Perceiving geometry: Geometric illusions explained by natural scene statistics*. New York: Springer.

Kanizsa, G. (1979). *Organization in vision: Essays on Gestalt perception*. Westport, CT: Praeger.

Konig, A., & Van Lier, R., (2007). 3D processing in the Poggendorff illusion. *Acta Psychologica*, *126*, 46–58.

Mack, A., Heuer, F., Villardi, K., & Chambers, D. (1985). The dissociation of position & extent in Müller-Lyer figures. *Perception & Psychophysics, 37*(4), 335–344.

MacLeod, D. I. A., & Willin, J. D. (1995). Is there a visual space? In R. D. Luce, M. D'Zmura, M. Hoffman, G. J. Iverson, A. K. Romney, & N. J. Markwah (Eds.), *Geometric representations of perceptual phenomena; Paper in honor of Tarow Indow* (pp. 47–60). Hillsdale, NJ: Lawrence Erlbaum.

Morgan, M. J. (1995). Visual illusions. In V. Bruce (Ed.), *Unsolved mysteries of the mind* (pp. 29–57). Mahwah, NJ: Lawrence Erlbaum.

Morgan, M. J. (1999). The Poggendorff illusion: A bias in the estimation of the orientation of virtual lines by second-stage filters. *Vision Research*, *39*(14), 2361–2380.

Morgan, M. J., & Casco, C. (1990). Spatial filtering & spatial primitives in early vision: An explanation of the Zöllner-Judd class of geometrical illusion. *Proceedings of the Royal Society of London B: Biological Sciences, 242*(1303), 1–10.

Morgan, M. J., Hole, G. J., & Glennerster, A. (1990). Biases and sensitivities in geometrical illusions. *Vision Research*, *30*(12), 1793–1810.

Potetz, B., & Lee, T. S. (2011). Scene statistics and 3D surface perception. In C W. Tyler (Ed.), *Computer vision: From surfaces to 3D objects* (pp. 1–26). New York: Chapman & Hall.

Robinson, J. O. (1998). *The psychology of visual illusion*. Mineola. NY: Dover. (Original work published 1972)

Rock, I. (1975). *An introduction to perception*. New York: Macmillan.

Sedgwick, H. A. (1986). Space perception. In K. Boff & J. Thomas (Eds.), *Handbook of perception and human performance* (pp. 21-2 to 21-49). New York: Wiley.

Smeets, J. B. J., Brenner, E., de Grave, D. D. J., & Cuijpers, R. H. (2002). Illusions in action: Some consequences of inconsistent processing of spatial attributes. *Experimental Brain Research*, *147*, 135–144.

Spehar, B., & Gillam, B. (2002). Modal completion in the Poggendorff illusion. *Psychological Science*, *13*(4), 306–312.

Tausch, R. (1954). Optische Taüschungen als artifizielle Effekte der Gestaltungs-prozesse von Grössen und Formenkonstanz in der natürlichen Raumwahrnchmung. *Psychologische Forschung, 24*, 299–348.

Warren, R. M., & Bashford, J. A. (1977). Müller-Lyer illusions: Their origin in processes facilitating object recognition. *Perception*, *6*(6), 615–626.

Worrall, N., & Firth, D. (1974). The components of the standard and reverse Müller-Lyer illusions. *Quarterly Journal of Experimental Psychology, 26*, 342–354.

Chapter 5
Visual Illusions in Action

Nicola Bruno

INTRODUCTION

Visual illusions such as those in this volume are robust effects. Some are strong, some are more subtle, but all are perceived readily. Now to "perceive" something means to report it in words or by means of suitable operations, such as matching or comparing. But what about acting? Will actions be affected by illusions that we perceive? For instance, consider the Ebbinghaus-Titchener illusion, which consists of two identical disks surrounded by larger and smaller flanker disks. When asked if these disks appear to have the same size, observers will report that they perceive the disk surrounded by larger flankers as smaller and the disk surrounded by smaller flankers as larger. This is the illusion: perceived size does not match physical size. But suppose that the disks were actual short cylinders and that the observers do not have to report what they perceive. Instead, they simply have to pick up one of the disks using the thumb and index of their preferred hand. To perform the action, they will quickly transport the hand forward while opening the fingers; at some point they will reach a maximum-aperture (MA) finger configuration, the hand will slow down for contact, and they will pick up the disk (Jeannerod, 1981). From studies in motor control (Marteniuk, Leavitt, MacKenzie, & Athenes, 1990) we know that, under ordinary circumstances, the MA correlates well with physical size. But the motor system has no way of knowing about physical size in advance. To modulate finger aperture, it must rely on visual information. Under ordinary circumstances, the processing of such information leads to fairly accurate representations of object size, but with the Ebbinghaus-Titchener display it does not. We should expect, therefore, that the illusion should show up in the MA just as it did in the verbal reports: observers-actors should open their fingers more when grasping the cylinder surrounded by smaller flankers and less so in the opposite situation.

This prediction seems so obvious that it hardly seems worth testing. Still, this prediction could be wrong, at least in certain conditions. In this chapter I present a review of the evidence supporting the notion that motor responses may resist visual illusions. The review will be selective, as dictated by space constraints, and highly critical because the interpretation of the relevant evidence remains controversial (for reviews presenting different positions on this issue, see Bruno, 2001; Cardoso-Leite & Gorea, 2010; Franz, 2001; Smeets & Brenner, 2006). The conclusions underscore the theoretical relevance, as well as the heuristic value, of studies of illusion effects on action responses. At the same time, however, they offer a warning concerning past interpretations of this literature and suggest that the proposed immunity of actions from visual illusions may be much more limited than proposed in earlier accounts.

INITIAL EVIDENCE

The first test of the effect of the Ebbinghaus-Titchener illusion on the in-flight MA of the fingers was performed by Aglioti, DeSouza, and Goodale (1995). In a clever experiment, these investigators created two versions of the illusion. In one, the targets were physically identical but perceived as different, whereas in the other they were physically different but perceived as identical. This second version was made possible by using a slightly larger target in the context of the big flankers and a slightly smaller one in the context of the small flankers. The two versions were alternated randomly, and participants were requested to pick up the disk on the left (right in other sessions) if they thought the disks were the same or to pick up the disk on the right (left) if they thought they were different. In all trials, MAs were recorded by using a video-based system that tracked light-emitting diodes attached to the thumb and index fingers and to the wrist. Results confirmed that participants were sensitive to the illusion, in that their choices of disks to pick up matched the two versions that had been created. Remarkably, however, the MAs recorded during the grasping movements did not reflect the perceptual differences but the physical sizes of the disks. On average, participants opened their fingers more with the larger disks and less with the smaller ones, irrespective of whether they perceived them as the same or different.

The authors interpreted this as demonstrating that the visual representation of object size that is used to drive grasping is different from the representation that forms the basis for verbal reports and recognition. The possibility that dissociable visual representations might be revealed in healthy participants in the context of visual illusions stirred a lot of interest, as it provided a new potential line of evidence in support of a specific

model of the human visual system. In this model (Milner & Goodale, 1995), the anatomical split that characterizes cortical projections after the primary visual areas is interpreted as the substrate of two independent subsystems: vision-for-perception (the ventral pathway from the primary visual cortex to the temporal lobe) and vision-for-action (the dorsal pathway to the posterior parietal lobe). On the basis of neuropsychological, neurophysiological, and imaging evidence, this "two visual systems hypothesis" (TVSH, often also called the "perception-action hypothesis"), explicitly proposed that vision-for-perception operates on object-relative, context-sensitive spatial representations. Coding the features of an object in relation to other objects in a scene is often useful to achieve invariance over changes of viewpoint, a critical step in object recognition. Vision for action, conversely, should privilege observer-relative, context-insensitive representations. When acting toward an object, the critical spatial relation is not with other objects in the environment but with one's own body (i.e., one's hand). That the size representation used to guide grasping would not be affected by the size of contextual flankers seemed therefore to be exactly what the hypothesis predicted.

INITIAL CRITICISMS

The interpretation of the results reported by Aglioti and colleagues (1995) has been criticized on several grounds. One early criticism noted that in the Aglioti et al. paradigm the perceptual and the motor tasks were not well matched in terms of attentional demands (Franz, Gegenfurtner, Bulthoff, & Fahle, 2000; Pavani, Boscagli, Benvenuti, Rabuffetti, & Farne, 1999). This is a fair point: when participants evaluated whether the disks were the same or different, they had to move the attentional focus from one disk to the other to perform the comparison. When they then performed the grasp, they needed to pay attention only to the disk on the side corresponding to the selected response. Thus, the observed size-contrast effect may have been determined by both contexts in the perception phase of the response but only by one in the action phase. This may have artifactually reduced the size-contrast effect in the motor task, making it harder to detect a significant difference. However, several later studies employed procedures whereby the perceptual and motor tasks were matched well but continued to find differences between the perception and motor task. Several studies (e.g., Daprati & Gentilucci, 1997; Haffenden & Goodale, 1988) used a modified perceptual task whereby participants reported perceived sizes by adjusting the distance between the finger and thumb without moving the hand forward. This form of cross-modal matching is sometimes called "manual estimation" (ME), and it involves the same attentional distribution as the motor response. These studies consistently reported larger effects of the illusion on the ME in comparison to the MA (for a review, see Franz, 2003). Results such as these rule out inappropriate matching of conditions. However, other criticisms have been raised that may continue to apply even to these better-designed studies. One particularly important point was raised by Franz, Fahle, Bulthoff, and Gegenfurtner (2001), who noted that the MA and the ME are not directly comparable because these measures behave quite differently as physical size is changed. The MA during a grasp reflects hand shaping in preparation for the final, slower approach until the fingers eventually touch the object. For this reason, the MA is always much larger than the physical size of the object, whereas its scaling as a function of physical size is compressed. In more precise terms, this means that the linear function relating MA and size has a nonzero intercept and a less-than-unitary slope (about 0.8 by most estimates; see Smeets & Brenner, 1999). The ME response, conversely, is a cross-modal match with regard to a spatial extent, a stimulus dimension that is generally believed to yield linear functions with zero intercept and unitary slope (Stevens, 1957). In fact, most studies found MA slopes greater than 1, and one found a slope as large as 1.6 (Franz, 2003).

Given these large differences in the scaling to physical size, it is problematic to measure the effect of the illusion on the ME and the MA by simply computing the raw differences between the large and the small flanker conditions. To see why, suppose we conducted a study and found that these raw differences were 1.6 mm in the ME data but only 0.8 mm in the MA data. When considering these raw effects, we might be led to conclude that the effect of the illusion on the ME was twice as large as that on the MA—a demonstration that the action resisted the illusion. However, the difference would depend entirely on the scaling of the two measures to physical size. In the ME, a difference of 1.6 mm could well correspond to a change in physical size of 1 mm, which is exactly the same physical size change that corresponds to a difference of 0.8 mm in the MA. To address this problem, Franz et al. (2001) proposed that the raw illusion effects should be corrected to make them comparable and argued that an appropriate correction consists in dividing the raw effects by the slopes derived from regressing against physical size. This correction expresses the MA and the ME in terms of equivalent physical changes, defining a common unit for both. When this correction is performed, many results no longer show a difference between the ME and the MA (Franz et al., 2001; see also Franz, 2001).

The importance of using corrected illusion effects when comparing across perceptual and motor responses is now generally recognized (for a reply see Schenk, Franz, & Bruno, 2011). Even after applying the appropriate corrections, however, there are still some studies that demonstrate differences between the ME and the MA. In one such study, for instance, Ganel, Tanzer, and Goodale (2008) employed a modification of the Ponzo illusion, which is usually demonstrated by two identical segments within a linear perspective context that causes one to appear farther away (and bigger) than the other, which appears closer (and smaller). In their modification, one segment (A) appeared clearly larger than another segment (B) just as in the standard illusion. The physical size of A, however, was in fact less, not more than that of B. Given the nature of this modified illusion display, Ganel et al. predicted that the MEs (perception) would show one pattern ($A > B$), whereas the MAs (action) would show the opposite ($A < B$), a double dissociation between perception and action. Their results confirmed this prediction. This pattern of results could not be altered by a change of scale, as this could not change the sign of the data.

In summary, early criticisms to Aglioti et al. (1995) played an important role in identifying methodological issues that had been previously overlooked and in raising the standards for performing comparisons between perceptual and motor responses in the context of illusions. While in general these initial criticisms failed to provide alternative explanations that could account for all published results without assuming separate vision-for-action and vision-for-perception subsystems, the ensuing debate identified still other issues. Among these, an important role has been played by issues related to the availability of visual information at different temporal phases of the action.

TIMING ISSUES

The visual control of actions evolves in time. After a preparatory phase, which uses the available visual and proprioceptive information to program the action, movement begins and new information becomes progressively available for online control. Since the pioneering work of Woodsworth (1899), it has been generally recognized that relatively simple actions such as reaching and grasping can be divided into an initial, ballistic phase and a slower control phase. The initial phase is believed to reflect the initial program, whereas the control phase uses feedback loops to refine and conclude the movement. The role of these different phases is traditionally investigated by techniques that remove visual feedback at different times during the action. For instance, the action can be performed under full visual guidance until its completion. In these conditions ("closed-loop" control), the kinematics of the action reflects contributions from both the preparatory program and online control loops. Alternatively, experimental procedures can remove visual feedback (e.g., by using glasses that can be rapidly switched from transparent to opaque) as soon as movement begins. In these conditions ("open-loop" control), the kinematics can reflect only the contribution of the preparatory phase. In addition to this basic distinction, different delays can be imposed between the removal of visual information and the initiation of the action. For instance, a go signal can be presented 3 or 5 s after the visual stimulus has been turned off. In these conditions, the action must be driven based on memorized rather than currently available visual information. An especially interesting limiting case for delayed control is the situation whereby the go signal actually coincides with the removal of visual information. In these conditions, there is only a small delay, which depends on the actor's reaction time. However, the action must still be driven based on the recent memory of the stimulus. This limiting case is therefore critically different from open-loop control, where the programming of the action is guaranteed to be complete before visual information is removed.

Several studies (e.g., Hu & Goodale, 2000) have shown that under delayed control, the effect of illusions on actions is comparable to the effect on perceptual measures such as the ME. This finding is consistent with another feature of the TVSH set forth by Milner and Goodale (1995). According to the TVSH, the vision-for-perception subsystem is slow and operates on longer-lasting representations, whereas the vision-for-action system is faster and employs short-lived representations. This difference is assumed to stem from the different functions of the two systems. Vision for perception is geared to producing conscious awareness of a stable environment. Vision for action, conversely, is interested only in providing information about the present relationship between the acting body and the object, a relationship that is bound to become obsolete as soon as movement begins. One would therefore predict that delayed actions would not be guided by vision-for-action, where the information soon decays after the visual stimulus is extinguished, but would require a switch to representations produced by vision-for-perception. Interestingly, even a relatively small delay as the mere actor reaction time seems to be sufficient to cause such a switch (Westwood & Goodale, 2003). These effects are not controversial, having been reported both by groups working within the TVSH and by groups that have been critical of Milner and Goodale's hypothesis (e.g., Franz, Hesse, & Kollath, 2009; Gentilucci, Chieffi, Daprati, Saetti, & Toni, 1996; Westwood, Heath, & Roy, 2000). However, they are of limited interest with regard to the interpretation of possible perception–action dissociations as they identify cases where no dissociation would be expected by the TVSH as well as by its critics.

More relevant to the issue of illusion effects on action are comparisons of effects on closed- and open-loop actions. Recall that the critical difference between these is that, in open-loop conditions, movement is driven by the preparatory program and cannot use online visual feedback. In closed-loop conditions, conversely, online feedback can be used to refine the movement after its ballistic onset. That such online corrections do take place in grasping has been documented by several studies (Castiello, Bennett, & Stelmach, 1993; Paulignan, Jeannerod, MacKenzie, & Marteniuk, 1991). Note that online corrections do not require internal models of object features but can operate simply on the basis of local error-reducing feedback loops. Because the initial studies based on the Aglioti et al. (1995) paradigm investigated closed-loop motor responses, it could be that the smaller effect of the illusion on these responses depended on online corrections during the grasp rather than on a context-insensitive visuomotor representation, which is produced in parallel with the context-sensitive representation for conscious recognition (Post & Welch, 1996; for a related conceptualization of this idea see also Glover, 2002). This alternative interpretation is attractive for at least two reasons. First, it preserves the functional distinction between vision-for-action and vision-for-perception, although it denies that two independent spatial representations are produced by these two subsystems and reduces its scope. Second, it is parsimonious.

However, whether this alternative explanation can account for the results of all studies of illusions and actions remains strongly controversial. In an early study, Westwood and Goddale (2003) measured the MA when grasping rectangular objects surrounded by larger or smaller flankers. In the critical comparison, vision was removed either at movement onset (open-loop) or at the presentation of a go signal (short delay). They reported that the size-contrast illusion affected the MA in the delay condition but not in the open-loop condition. However, this study did not compute corrected effects, which may change the pattern of results as the slope of open-loop and delayed MAs were markedly different. Conversely, more recent meta-analyses on grasping

the Ebbinghaus-Titchener illlusion (Franz & Gegenfurtner, 2008) and on the well-known Müller-Lyer illusion (Bruno & Franz, 2009) did use corrected effects while combining large groups of studies. These two papers converged in showing that grasping is less affected by illusions than perception if performed under full visual guidance (closed-loop) but is affected by the illusion as much as perception (or delayed actions) whenever the grasp is performed without visual feedback (open-loop). The double-dissociation paper by Ganel et al. (2008) also included a control open-loop condition that replicated the findings of the main closed-loop experiment. As we have seen, this result would not be changed by computing corrected effects. However, it has been pointed out that the methodology of the open-loop condition in this study was not completely adequate to insure that there was no visual feedback (Schenk, Franz & Bruno, 2011). Finally, Franz et al. (2009) performed a careful study of the MA when grasping the Müller-Lyer illusion. They systematically manipulated the availability of visual feedback by removing vision never, after two-thirds or one-third of the movement, immediately at movement onset, at a go-signal, or after a 5 s delay. They observed a gradual decrease of corrected illusion effects as the availability of feedback increased, with no evidence of a shift between dorsal and ventral control. Taken together, these studies suggest that although the issue is not settled, the possibility that actions resist illusion only when visual feedback can be used for online correction should be taken quite seriously.

OTHER ACTIONS

Proponents of the TVSH have mostly limited the development of the theory to the visual control of grasping. However, the debate on the hypothesized immunity of actions from illusions has stimulated renewed interest toward earlier reports of reduced effects of illusions on ballistic pointing (Mack, Heuer, Villardi, & Chambers, 1985) as well as a production of a large body of research on the effects of various illusions on other actions, such as, again, pointing (e.g., Bruno & Bernardis, 2003; de Grave, Brenner, & Smeets, 2004; Gentilucci et al., 1996), bimanual reaching (e.g., Bruno & Bernardis, 2002), saccading (e.g., Bernardis et al., 2005; Bruno et al., 2010; de Grave & Bruno, 2010; Knox & Bruno, 2007), flicking (Hartung, Schrater, Bulthoff, Kersten, & Franz, 2005; Kroliczak, Heard, Goodale, & Gregory, 2006), and even stepping (McCarville & Westwood, 2006; Wraga, Creem, & Proffitt, 2000) and hopping (Glover & Dixon, 2004). The evidence from these studies is mixed, and a comprehensive account of this large body of literature is currently lacking. In particular, a largely unresolved issue is whether data from different actions can be meaningfully compared with those from grasping and, if so, how. Especially relevant with this regard are the results of a recent meta-analysis of several studies of pointing on the Müller-Lyer illusion (Bruno et al., 2008). This meta-analysis demonstrated robust corrected effects of the illusion on pointing in delayed conditions and almost null effects both in closed-loop and in open-loop conditions. These results therefore are more in line with the predictions of the TVSH and stand in contrast to those of comparable meta-analyses on grasping (see previous section). This in turn might point to a fundamental difference between the neural control of pointing and grasping, or it might depend on some other factor that is not yet understood.

CONCLUSIONS

About 20 years ago, a series of early reports suggested that actions may resist visual illusions. This proposal has attracted a lot attention, not only for its implications for functional models of the visual system but also for its potential to resolve longstanding controversies between direct and constructivist theories of perception (see Norman, 2002). This proposal has been also remarkably heuristic, stimulating the production of a large body of research on perception and action. This proposal remains, however, deeply controversial. Despite a number of attempts to resolve the issue, there is still no consensus view as to whether perception and action can be dissociated in the context of illusions and under what conditions. Based on the available evidence, this chapter has suggested that the weight of the evidence seems to lean in favor of an interpretation in terms of visual feedback. Actions resist visual illusions when local, online error correction processes can refine the initial, context-sensitive representation that drives both perception and action-related motor programs. This interpretation preserves the functional distinction between vision-for-perception and vision-for-action but greatly reduces the scope of the latter in comparison with the initial proposal of Milner and Goodale (1995).

REFERENCES

Aglioti, S., DeSouza, J. F., & Goodale, M. A. (1995). Size-contrast illusions deceive the eye but not the hand. *Current Biology*, *5*, 679–685.

Bernardis, P., Knox, P. C., & Bruno, N. (2005). How does action resist visual illusion? Uncorrected oculomotor information does not account for accurate pointing in peripersonal space. *Experimental Brain Research*, *162*, 133–144.

Bruno, N. (2001). When does action resist visual illusions? *Trends in Cognitive Sciences*, *5*(9), 379–382.

Bruno, N., & Bernardis, P. (2002). Dissociating perception and action in Kanizsa's compression illusion. *Psychonomic Bulletin & Review*, *9*, 723–730.

Bruno, N., & Bernardis, P. (2003). When does action resist visual illusions? Effector position modulates illusory influences on motor responses. *Experimental Brain Research*, *151*(2), 225–237.

Bruno, N., Bernardis, P., & Gentilucci, M. (2008). Visually guided pointing, the Muller-Lyer illusion, and the functional interpretation of the dorsal-ventral split: Conclusions from 33 independent studies. *Neurosci Biobehav Rev*, *32*(3), 423–437.

Bruno, N., & Franz, V. H. (2009). When is grasping affected by the Müller-Lyer illusion? A quantitative review. *Neuropsychologia*, *47*(6), 1421–1433.

Bruno, N., Knox, P., & de Grave, D. (2010). A metanalysis of the effect of the Müller-Lyer illusion on saccadic eye movements: No general support for a dissociation between perception and oculomotor action. *Vision Research*, *50*, 2671–2682.

Cardoso-Leite, P., & Gorea, A. (2010). On the perceptual/motor dissociation: A review of concepts, theory, experimental paradigms and data interpretations. *Seeing and Perceiving*, *23*(2), 89–151.

Castiello, U., Bennett, K., & Stelmach, G. (1993). Reach to grasp: The natural response to perturbation of object size. *Experimental Brain Research*, *94*, 163–178.

Daprati, E., & Gentilucci, M. (1997). Grasping an illusion. *Neuropsychologia*, *35*(12), 1577–1582.

de Grave, D. D. J., Brenner, E., & Smeets, J. B. J. (2004). Illusions as a tool to study the coding of pointing movements. *Experimental Brain Research*, *155*(1), 56–62.

de Grave, D. D. J., & Bruno, N. (2010). The effect of the Müller-Lyer illusion on saccades is modulated by spatial predictability and saccadic latency. *Experimental Brain Research*, *203*, 671–679.

Franz, V. H. (2001). Action does not resist visual illusions. *Trends in Cognitive Sciences*, *5*(11), 457–459.

Franz, V. (2003). Manual size estimation: a neuropsychological measure of perception? *Experimental Brain Research*, *151*, 471–477.

Franz, V. H., Fahle, M., Bulthoff, H. H., & Gegenfurtner, K. R. (2001). Effects of visual illusions on grasping. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(5), 1124–1144.

Franz, V. H., & Gegenfurtner, K. R. (2008). Grasping visual illusions: Consistent data and no dissociation. *Cognitive Neuropsychology*, *25*(7–8), 920–950.

Franz, V. H., Gegenfurtner, K. R., Bulthoff, H. H., & Fahle, M. (2000). Grasping visual illusions: No evidence for a dissociation between perception and action. *Psychological Science, 11*(1), 20–25.

Franz, V. H., Hesse, C., & Kollath, S. (2009). Visual illusions, delayed grasping, and memory: No shift from dorsal to ventral control. *Neuropsychologia*, *47*(6), 1518–1531.

Ganel, T., Tanzer, M., & Goodale, M. A. (2008). A double dissociation between action and perception in the context of visual illusions: Opposite effects of real and illusory size. *Psychological Science, 19*(3), 221–225.

Gentilucci, M., Chieffi, S., Daprati, E., Saetti, M. C., & Toni, I. (1996). Visual illusion and action. *Neuropsychologia*, *34*(5), 369–376.

Glover, S. (2002). Visual illusions affect planning but not control. *Trends in Cognitive Sciences*, *6*, 288–292.

Glover, S., & Dixon, P. (2004). A step and a hop on the Müller-Lyer: Illusion effects on lower-limb movements. *Experimental Brain Research*, *154*(4), 504–512.

Haffenden, A. M., & Goodale, M. A. (1998). The effect of pictorial illusion on prehension and perception. *Journal of Cognitive Neuroscience*, *10*(1), 122–136.

Hartung, B., Schrater, P. R., Bulthoff, H. H., Kersten, D., & Franz, V. H. (2005). Is prior knowledge of object geometry used in visually guided reaching? *Journal of Visualization*, *5*(6), 504–514.

Hu, Y., & Goodale, M. A. (2000). Grasping after a delay shifts size-scaling from absolute to relative metrics. *Journal of Cognitive Neuroscience*, *12*(5), 856–868.

Knox, P., & Bruno, N. (2007). When does action resist visual illusion? The effect of the Müller-Lyer illusion on reflexive and voluntary saccades. *Experimental Brain Research*, *181*, 277–287.

Kroliczak, G., Heard, P., Goodale, M. A., & Gregory, R. L. (2006). Dissociation of perception and action unmasked by the hollow-face illusion. *Brain Research*, *1080*(1), 9–16.

Jeannerod, M. (1981). Intersegmental coordination during reaching at natural visual objects. In J. Long, & A. Baddeley (Eds.), *Attention and performance* (Vol. 11, pp. 153–168). Hillsdale, NJ: Lawrence Erlbaum.

Mack, A., Heuer, F., Villardi, K., & Chambers, D. (1985). The dissociation of position and extent in Müller-Lyer figures. *Perception & Psychophysics*, *37*(4), 335–344.

Marteniuk, R. G., Leavitt, J. L., MacKenzie, C. L., & Athenes, S. (1990). Functional relationship between grasp and transport components in a prehension task. *Human Movement Science, 9*, 149–176.

McCarville, E. M., & Westwood, D. A. (2006). The visual control of stepping operates in real time: Evidence from a pictorial illusion. *Experimental Brain Research*, *171*(3), 405–410.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action.* Oxford: Oxford University Press.

Norman, J. (2002). Two visual systems and two theories of perception: An attempt to reconcile the constructivist and ecological approaches. *Behavioral and Brain Sciences, 25*, 73–96.

Pavani, F., Boscagli, I., Benvenuti, F., Rabuffetti, M., & Farne, A. (1999). Are perception and action affected differently by the Titchener circles illusion? *Experimental Brain Research*, *127*(1), 95–101.

Paulignan, Y., Jeannerod, M., MacKenzie, C., & Marteniuk, R. (1991).Selective perturbation of visual input during prehension movements: II. The effects of changing object size. *Experimental Brain Research*, *87*, 407–420.

Post, R. B., & Welch, R. B. (1996). Is there dissociation of perceptual and motor responses to figural illusions? *Perception*, *25*(5), 569–581.

Schenk, T., Franz, V., & Bruno, N. (2011). Vision-for-perception and vision-for-action: Which model is compatible with the available psychophysical and neuropsychological data? *Vision Research*, *5*, 812–818.

Smeets, J. B., & Brenner, E. (1999). A new view on grasping. *Motor Control*, *3*(3), 237–271.

Smeets, J. B., & Brenner, E. (2006). 10 years of illusions. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(6), 1501–1504.

Stevens, S. S. (1957). On the psychophysical law. *Psychological Review, 64*(3), 153–181.

Westwood, D. A., & Goodale, M. A. (2003). Perceptual illusion and the real-time control of action. *Spatial Vision*, *16*(3–4), 243–254.

Westwood, D. A., Heath, M., & Roy, E. A. (2000). The effect of a pictorial illusion on closed-loop and open-loop prehension. *Experimental Brain Research*, *134*(4), 456–463.

Woodworth, R. (1899). The accuracy of voluntary movement. *Psychological Review Monograph, 3*(2), 1–114.

Wraga, M., Creem, S. H., & Proffitt, D. R. (2000). Perception-action dissociations of a walkable Müller-Lyer configuration. *Psychological Science*, *11*(3), 239–243.

Chapter 6
Motion Illusions in Man and Machine

Cornelia Fermüller

INTRODUCTION: THEORIES OF ILLUSIONS

Optical illusions have been studied in visual psychology since its early history, and the formation of theories followed closely. Since then there has been an enduring interest in the subject. Theorists of perception have looked at illusions as test instruments for theory, an effort that originated from the founders of the Gestalt school. An important strategy in finding out how correct perception operates is to observe situations in which misperception occurs. Any theory, to be influential, must be consistent with the facts of correct perception but also must be capable of predicting the failures of the perceptual system. While in the past only psychologists studied illusions, in recent decades they have been joined by scientists of other mind-related fields such as neurology, physiology, philosophy, and the computational sciences, examining the problem from different viewpoints with the use of different tools (Gillam, 1998; Palmer, 1999).

It is clear that vision cannot be veridical, or, as Palmer (1999, p. 7) puts it, vision cannot give us "a clear window into reality." Visual perception is very complex as it includes a large set of processes that start at the formation of images on our eyes and provide us with an interpretation of the objects and events in the scene. For the purpose of discussion, we distinguish three kinds of processes: processes of image formation, processes of building geometric models of the spatiotemporal scene, and processes of interpreting the scene in terms of the concepts we know. According to this division, we can classify illusions of vision into *physiological illusions* due to the image formation or very early representation of images; *optical illusions in the classic sense*, which are about perceiving nonveridical surfaces and motions; and *cognitive illusions* resulting from an ambiguous interpretation.

It does not surprise us that there are limitations to the early processes. Our eyes can be likened to an optical apparatus that collects the light emitted from surfaces in the world, and, like any measuring device, the eye has its limitations. It is limited by the spatial and temporal resolution of the photoreceptors and their adaptive sensitivity to light. These limitations can account for many of the *physiological illusions*; for example afterimages, the apparent motion effect, or the illusion of a fast-rotating object appearing to rotate in opposite direction of its true movement (the wagon-wheel effect). It is also no surprise that our interpretation of scenes is not unique. Cognitive interpretation requires a high level of knowledge and memory. This knowledge is acquired from our experiences and thus can vary between individuals. *Cognitive illusions* may help us understand the attention mechanisms involved in complex interpretation and how our mind organizes high-level visual information. Examples of such illusions are ambiguous figures, such as the old–young woman illusion, or paradox illusions such as the impossible staircase.

The illusions most studied by vision scientists are the *optical illusions in the classic sense*, which are due to the processes of making models of the scene or, as computational scientists say, the processes of "reconstructing the scene." Illusions in this category include the geometric-optic illusions, illusions of lightness and color, visual completion illusions, and illusions of motion. One not familiar with the study of vision may think that reconstruction is simple, but clearly it is not. The problem often is summarized as "the vision system has to solve the inverse optics problem." The images formed on our retina are projections of the three-dimensional (3D) world on two-dimensional (2D) imaging surfaces, and it is up to the vision processes to reconstruct the third dimension. The light incident on the retina depends on many factors, such as the different light sources, the properties of the reflecting surfaces, and interreflections between objects, and it is up to the vision processes to separate the different lighting components. The image intensity pattern in a single image may not carry enough information to divide the image into different surfaces and objects, and the vision system must use the best information available to do so. In these processes of recovering the surfaces and the movements of the scene from the 2D light patterns, things can go wrong and misperception occurs, giving us a window into the mechanisms underlying the reconstruction.

Most works on computational modeling have been influenced by the computational theory of Marr (1982). This theory describes vision as information processing, which must be understood at three levels: at the computational level (What does the system do and why does it do it?), at the algorithmic or representational level (What are the processes and representations?), and at the physical level

(What is the physical realization [i.e., the hardware] or the neural structure of biological systems?) Marr suggests that what the system computes is 3D descriptions of the world from 2D images, and this is algorithmically implemented in a sequence of stages. First, local features, such as line features, feature points, and local image motion are extracted; these are linked to obtain contours and connected regions, then the shape of the surfaces as seen from a given viewpoint is derived from cues such as motion, texture, or shading, and finally the scene models from multiple views are combined into a single 3D model.

Most computational theories of optical illusions seek explanations either at the algorithmic level or the physical level. For example, theories of lightness illusions and geometric optical illusions predominantly refer to the representations, usually the quantities at the first stage of the Marr representation, such as edges, lines, junction corners, and regions, and the mechanisms of computing them (e.g., Grossberg & Todorovic, 1988; Morrone & Burr, 1988; Neumann & Mingolla, 2003), and many theories of motion illusions refer to the hardware mechanisms implementing the 2D image motion (e.g., Conway, Kitaoka, Yazdanbakhsh, Pack, & Livingstone, 2005; Mingolla, 2003). While these theories seek specific mechanisms, which are hypothesized to exist in the vision system, the theories proposed here are at a higher level of abstraction. We argue the reason for the malfunctioning of image motion estimation is to be found at the level of the abstract computational model, no matter how the computation is algorithmically or physically implemented.

At the computational level, many theories of optical illusions do not consider the process of obtaining world models from images. Since many optical illusions are patterns on a 2D surface, it is often assumed that the goal of vision is to create an accurate image of these 2D dimensional patterns. This applies for most classical theories. A notable exception is Gregory's theory (Gregory, 1968) of geometric-optical illusions, which supposes that 2D patterns are interpreted as flat projections of 3D scenes. In the case of motion, reference to a 3D interpretation is imminent. It is clear that motion is a cue to space-time geometry. The goal of early motion interpretation is to build spatiotemporal models of the scene, and considering the 3D interpretation as the goal of motion processing, as explained next, can explain why independent motions are seen in static images.

MOTION PERCEPTION

We can see the motion in the world because our vision has mechanisms to perceive changes in the image signal over time. It has been proposed that the human vision system has multiple mechanisms (Chubb & Sperling, 1988), some tuned to instantaneous local motion, some to regional motion, and some to large motion of segmented parts. The prominent mechanism, which allows us to see continuous motion, is based on computing the instantaneous changes of the incoming light. The vision system records the changes in the incoming light on the retina and then relates them in further processes to the physical changes in the world.

There are limitations to this motion mechanism, simply because of physical limitations, or "hardware" constraints. It is limited in temporal resolution. For this reason motion perception can be induced with a sequence of static images. These effects are known as apparent motion (Exner, 1875; Wertheimer, 1912). When presented with video, we perceive a continuous motion and not a succession of 24 image frames. There are limitations to the velocity that can be perceived, as is evidenced by the wagon-wheel effect. Apparent motion can also be induced, simply by gradually changing the intensity in a pattern. The system then computes a temporal change in the intensity and perceives motion, the so-called reverse phi effect (Anstis, 1970; Anstis & Rogers, 1986). Furthermore, biological systems do not, as suggested by video technology, compute image motion directly from the difference between the frames in a video sequence. Instead, the neurons in the early stages of processing produce responses whenever the change of intensity at an image point exceeds some value (Laughlin, 1981; Srinivasan, Laughlin, & Dubs, 1982). This difference in response time can account for the difference in velocity perceived in motion sequences on patterns with significantly different contrast in different regions, such as the stepping-feet illusion (Anstis, 2003).

Visual motion analysis encompasses two distinct processes: the computation of image motion and the interpretation of the image motion in terms of the 3D scene and 3D movement. The early interpretation involves computing the observer's own motion, or ego-motion, the segmentation of the scene into surfaces at different depth, detection of objects that move in the scene, and the estimation of their movement. Computational approaches treat the whole process strictly in two successive steps, first computing image motion then interpreting it. The many feedback loops found from neural studies indicate that biological systems do not implement it in such a strict bottom-up fashion, but rather the processes are intertwined and solved iteratively. How these feedback loops work is still not clear. But this does not matter for our model here. It is important only to understand that image motion contributes to ego-motion and segmentation, but the three processes are not independent but rather shape each other.

IMAGE MOTION ESTIMATION AND ITS LIMITATIONS

Image motion is obtained from the changing image patterns on the retina over time. As a representation for image motion we use the concept of the *optical flow*. The optical flow vector associated with an image point represents the point's instantaneous movement. The corresponding field of all measurements (a vector at every image point), the so-called optical flow field, represents an approximation of the projection of the 3D scene points' movement on the image. Computational considerations as well as biological measurements suggest that optical flow is computed in two stages (Adelson & Movshon, 1982).

In a first stage, considering only local image measurements in a small spatiotemporal region, the velocity component perpendicular to linear features is computed. The situation is illustrated in Figure I.6-1a. The velocity vector of a line segment viewed through a small aperture is inherently ambiguous, as it is consistent with any vector falling

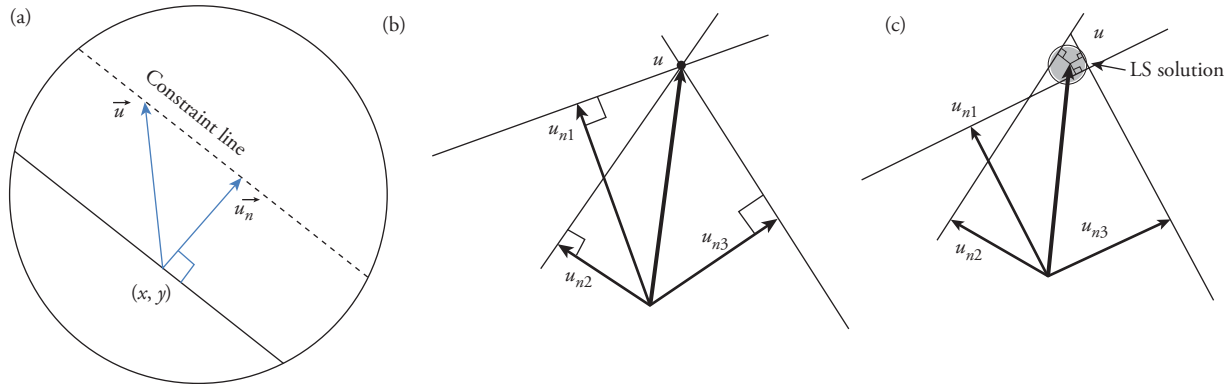


Figure I.6-1. (a) The aperture problem: Observing a line feature (solid line) through a small aperture, we cannot determine where exactly the point (*x, y*) moves to. We can only obtain a constraint line (dashed line), which provides the component of flow perpendicular to the line. (b) If there is no noise, the constraint lines of different image points with their different gradient directions will intersect in one point. (c) In the presence of noise, we estimate the optical flow by minimizing some distance; in the case of LS estimation this is the average distance to the constraint lines. (For an animated explanation, see www.cfar.umd.edu/~fer/optical/movement1.html)

on the constraint line (Wallach, 1935). Only the velocity component perpendicular to the line feature in the direction of the motion is well defined. This component of the image motion is referred to as *normal flow*, and the ambiguity is referred to as the *aperture problem* (Marr & Ullman, 1981). In order then to derive in a second stage the complete 2D optical flow, one-dimensional normal flow measurements from line features in different directions in a neighborhood need to be combined and the optic flow estimated from them.

When we compute optical flow, we treat the image intensity as a smoothly varying function of space and time, and we derive the optical flow from the changes in the image intensity. Either we use filters that approximate the derivatives in the image domain, or we use filters tuned to certain frequencies of space and time, such as biological plausible motion filters. It is generally considered that image motion can be estimated quite well at locations in the image where the assumptions hold; these are textured regions on smooth surfaces. It is also known that there are problems at the locations where the scene changes, that is, at the boundaries of 3D surfaces, because there the image intensity function is discontinuous. At these locations occlusions make it difficult to obtain accurate normal flow, and the estimation of optical flow is even more challenging (Horn & Schunck, 1981).

In this chapter we show that, even within areas of smooth flow (i.e., the regions corresponding to smooth-textured 3D surfaces), the computation of flow poses a problem, and both normal flow and optical flow are affected, as summarized next.

The estimation of normal flow is erroneous for certain signals, because real-time systems estimate temporal changes using causal filters. That is, motion is found from the response of spatiotemporal filters, with the filters tuned to spatial frequencies being symmetric but the filters tuned to temporal frequencies being asymmetric. This is because real-time systems receive as input data from the present and the past. The so-called causal filters used for processing this data are asymmetric with greater weight given to recent input than older input, because otherwise the processing would be much delayed. Such causal filters misestimate the image motion in asymmetric image signals for certain spatial frequencies.

The estimation of optical flow cannot be accurate because of statistical reasons. Like any real system, our vision has to deal with noisy data. The local normal flow measurements can be estimated only within some range of the actual value; in other words, they are noisy. Because of this noise, the estimation of optical flow using as input normal flow is statistically biased. Thus the expected value and the actual value are different. From a statistical viewpoint, correcting for the bias would require knowledge of the noise parameters. The noise parameters would have to be estimated from the data, but they are difficult to obtain, because the parameters are not static; instead they change in unpredictable and complex ways, as the scene, the viewing conditions, and the lighting changes. There always is bias in the estimation, but in most situations the bias is insignificant. It depends on the texture in the image and can be large for regions with a dominant gradient distribution.

The next two sections discuss the misestimation of image motion in more detail and how it can account for optical motion illusions. The next section discusses the challenges in the estimation of optical flow and the following section the challenges in normal flow estimation. The chapter concludes with a discussion on the role of early motion processes and their interaction with other processes of static cues.

BIAS IN OPTICAL FLOW ESTIMATION

There are two dominant models for the computation of optical flow in the literature: gradient-based methods and frequency-based methods. Both methods derive optical flow in a two-stage process: first information about the one-dimensional motion component of local edges or single spatial frequencies is obtained; then, the individual measurements within some neighborhood are combined into an estimate of optical flow. The two methods are faced with similar noise issues and can be given a similar mathematical modeling.[1]

The monochromatic light distribution on the retina can be described as a function *I(x, y,t)* that specifies the intensity at a point (*x, y*) at time *t*. Let us denote the instantaneous motion of a point in the image by a translation of velocity *v* with horizontal and vertical speed components (*u, v*). Assuming that the intensity at a point over a small time interval does not change, that is

$$I(x,y,t) = I(x-ut, y-vt, 0),$$

we obtain from a first-order Taylor expansion of this equation the following constraint relating the spatial and temporal derivatives I_x, I_y, and I_t and the image motion

$$I_x u + I_y v = -I_t. \tag{1}$$

This equation, called the *optical flow constraint equation* (Horn & Schunck, 1981), defines at a point the component of flow in the direction of the spatial gradient (I_x, I_y), or in geometric terms, it defines the component of the flow perpendicular to the edge (as illustrated in Fig. I.6-1a), and this component is termed the *normal flow*. Similarly, taking the 3D Fourier transform of equation 1, we obtain (Adelson & Bergen, 1985; Watson & Ahumada, 1985)

$$u\omega_x + v\omega_y = -\omega_t, \tag{2}$$

where ω_x, ω_y denote the spatial frequencies and ω_t the temporal frequency. We can also interpret this equation as follows: given the dominant spatial frequency at an image point, we know the flow component parallel to the direction of this frequency. The frequencies can be computed from data of a spatiotemporal volume using filters tuned to certain frequencies.

To derive the optical flow, a second mathematical constraint must be formulated. Most computational methods assume that the flow varies smoothly. The very simplest assumption is that the flow is constant within a neighborhood. This is called the *intersection of constraints* model. We use this model to keep the exposition simple.

Let us consider that in a neighborhood we have *n* measurements. If considering a gradient based approach, each measurement *i* provides an equation (eq. 1), and we obtain an overdetermined system of equations, which we write as

$$A\vec{u} = \vec{b}, \tag{3}$$

where A denotes the $n \times 2$ matrix of spatial gradients (I_{x_i}, I_{y_i}) and $\vec{b}$ the *n*-dimensional vector of temporal derivatives (I_{t_i}). Similarly, if considering a frequency-based approach (eq. 2), A and $\vec{b}$ are the spatial and temporal frequencies $(\omega_{x_i}, \omega_{y_i})$ and ω_{t_i}. Each equation in (3) provides a constraint line, and in principle, from two measurements we can solve for the optical flow. If all the measurements (I_{x_i}, I_{y_i}) and I_{t_i} were perfect, all constraint lines would intersect in a point. However, because they are noisy, we have to use an estimator to find a solution that is best with respect to some criterion. Standard least squares (LS) estimation finds the point that has minimum normal distance to all the constraint lines (see Fig. I.6-1b and I.6-1c). The LS solution of equation 3 for the flow yields the estimate

$$\vec{u} = (A^T A)^{-1} A^T \vec{b}, \tag{4}$$

where A^T and A^{-1} denote the transpose and the inverse of matrix A, respectively.

The observed data is always noisy, or we say the observations $A = (a_{1_i}, a_{2_i})$ and (b_i) are corrupted by errors. We use unprimed letters, primed letters, and δs to denote the estimates, actual values, and errors, respectively. We can say that the observations are composed of the true values (A', b') and the errors $(\delta A, \delta b)$, and rewrite equation 3 as

$$(A + \delta A)\vec{u} = (\vec{b} + \delta\vec{b}). \tag{5}$$

It is well known that LS estimation generally is biased (Fuller, 1997; van Huffel & Vandewalle, 1991). It provides an unbiased solution if noise occurs in variable *b* only, that is in our case the length of the normal flow, but it is biased if there is noise also in the explanatory variables *A*, in our case, the direction for the normal flow. To clarify: the concept of *statistical bias* means that if we had many, many noisy measurements to estimate the unknown parameter, then the expected value of our estimate (in the limit) would be different from the true value.

Under some simplifying assumptions (identical and independent random variable δ *A* and δ *b* with zero mean and variance σ^2), the estimate $E(\vec{u})$ and the true value $\vec{u}'$ are related as (Fuller, 1997)

$$E(\vec{u}) = \vec{u}' - \sigma^2 \left(\lim_{n\to\infty} \left(\frac{1}{n} A'^T A' \right) \right)^{-1} \vec{u}', \tag{6}$$

which implies that the estimate is asymptotically biased. The second term in equation 6 is the bias. It depends on the amount of noise, σ^2, the texture, which is described by matrix $M' = (A'^T A')$, and the relationship between the actual flow and the texture. An analysis of the bias term reveals that large variance in δ*A*, an ill-conditioned A', or a $\vec{u}'$, which is oriented close to the eigenvector of the smallest eigenvalue of $A'^T A'$, all could increase the bias and push the LS solution away from the real solution. Generally it leads to an underestimation of the parameters.

To understand why there is underestimation, consider the simpler problem of estimating the one-dimensional unknown *x* from a set of linear equations of the form $ax = b$. As an example, consider two measurements with a symmetric noise distribution. Let these measurements be $a_1 = a' + \delta a$, $a_2 = a' - \delta a$. (Noise symmetric in *b* does not have an effect on the bias, so we do not need to consider it in our example; thus $b_1 = b_2 = b'$). The exact solution of the equation $a'x = b'$ amounts to $x = \frac{b'}{a'}$. The LS solution of the given linear equations amounts to

$$x = \frac{\sum_{i=1}^{2} a_i b_i}{\sum_{i=1}^{2} a_i^2} \tag{7}$$

Substituting for the values of a_1, a_2, b_1, b_2, we obtain

$$x = \frac{(a' + \delta a + a' - \delta a)b'}{(a' + \delta a)^2 + (a' - \delta a)^2} = \frac{a'b'}{\left(a'^2 + 2(\delta a)^2\right)} < \frac{a'b'}{a'^2} = \frac{b'}{a'} \quad (8)$$

This intuitively explains the bias. Basically, the bias originates from the quadratic term a^{-2}, and it is due to the variance of δa. In the 2D case a^{-2} becomes $(A^T A)^{-1} = (A' + \delta A^T (A' + \delta A))^{-1}$ and can be obtained from a second-order Taylor expansion of equation 4.

The information about the bias is encoded in $M' = A^T A$, the matrix of spatial gradients, which describes the texture in a patch. In the case of a uniform distribution of the image gradients, the bias is solely in the length of the computed optical flow; there is an underestimation. Consider two dominant gradient directions. The signal to noise ratio is smaller in the direction of fewer measurements. Thus there is more underestimation in the direction of fewer measurements and less underestimation in the direction of more measurements. As a result, the estimated flow is biased downward in size and biased toward the major direction of the gradients in the patch.

Ouchi Illusion

The bias provides an explanation for many motion illusions, including the well-known Ouchi pattern. This striking illusory pattern created by the graphic artist H. Ouchi, and introduced as illusion to the research community by Spillmann, Heitger, and Schuller, (1986) consists of two black-and-white rectangular checkerboard patterns oriented in orthogonal directions—a background orientation surrounding an inner ring (Fig. I.6-2a). Small retinal motions, or slight movements of the paper, evince a segmentation of the inset pattern and motion of the inset relative to the surround. The illusion occurs for a variety of viewing distances and angles. Some observers report an apparent depth discontinuity, with the center floating as it moves above the background (Spillmann, Tulunay-Keesey, & Olson, 1993). In a nutshell, the cause of the illusion is differently biased flow vectors in the two patterns (Fermüller, Pless, & Aloimonos, 2000), as follows.

The tiles used to make up the pattern are longer than they are wide. Thus, in any small region there are many more normal flow measurements in one direction than the other. Since the tiles in the two regions of the figure have different orientations, the estimated regional optical flow vectors are different (they have different bias). The global motion of the pattern is computed from the data in the surround. When looking at the static pattern under free viewing conditions, small retinal eye movements cause the motion, which are estimated from the periphery. The difference between the bias in the inset and the bias in the surrounding region is interpreted as motion of the ring. An illustration of the estimated flow and the difference of the estimated and veridical flow is given in Figure I.6-2b and I.6-2c for the case of motion along the first meridian (to the right and up). In addition to computing flow, the visual system also performs segmentation, which is why a clear relative motion of the inset is seen.

The theory can account for variations in the pattern. In Figure I.6-2 the ratio of the two sides of the rectangle is 4:1; a smaller ratio is predicted to lead to smaller bias as has been perceived. The bias also depends on the direction of movement. It is predicted to be largest for a diagonal movement and smallest for movements parallel to the sides, as can also be perceived. Khang and Essock (1997) created variations of the Ouchi pattern by replacing the periodic rectangular with other periodic functions that are "smoother" and have more local gradient orientations. The bias model predicts that, with the spreading of directions, the amount of bias in the estimated flow decreases, which explains the decrease in the perceived illusory motion found in the experiments.

Many other illusory percepts can be explained by the bias, such as the erroneous perception of motion in plaid patterns, the Revolving Wheels Illusion (Pinna & Brelstaff, 2000), the perceived motion of contracting and expansion in rotating spirals, and the rigid versus no-rigid motion perceived in translating sinusoids of high and low amplitude, respectively (see the Companion Website www.cfar.umd.edu/~fer/optical/).

A Note on Estimation

We have shown that LS estimation is biased. One may ask: Is the vision system really doing LS? Could it do better? Looking at the statistical problem, LS estimation is only one of many possible estimators. The question, then, is: Are there other estimators that do not suffer from this problem? An elaborate discussion on the statistics of optical flow estimation can be found in Fermüller, Shulman, and Aloimonos (2001). The analysis there shows that all methods—gradient-based, energy-based, and correlation methods—suffer from bias, and for many statistical estimators the bias is an underestimation. The reason it is so difficult to correct for the bias is that this would require knowledge of the noise parameters. These parameters would need to be estimated from the data, but usually there is not enough data available about the value of the noise. It might be possible if the parameters were static, but instead they change in unpredictable ways with the change of environment, the lighting, and viewing conditions.

Statistical estimators must deal with two components: one is bias, and the other is variance; and there is a trade-off between the two. Generally an estimator correcting for bias increases the variance while decreasing the bias. Thus, if in some cases there is data to extract the statistics of the noise, for example when we look at the same static scene for some time, theoretically the best thing to do is to partially correct the bias. This does not change the form of the bias but decreases its amount. We believe that the human vision system, when it has access to enough data, follows this theoretically best approach. Two observations make us believe so. First, illusory perception in many optical illusions weakens after extended viewing, in particular when subjects are asked to fixate (Helmholtz, 1962; Yo & Wilson, 1992). In these cases the noise parameters stay fixed, and the vision system can reasonably estimate them well and partially correct. Second, in experiments (Ji & Fermüller, 2006) we varied the variance by changing the density of the texture in a pattern. Theoretically, a better correction and thus better estimation is possible for

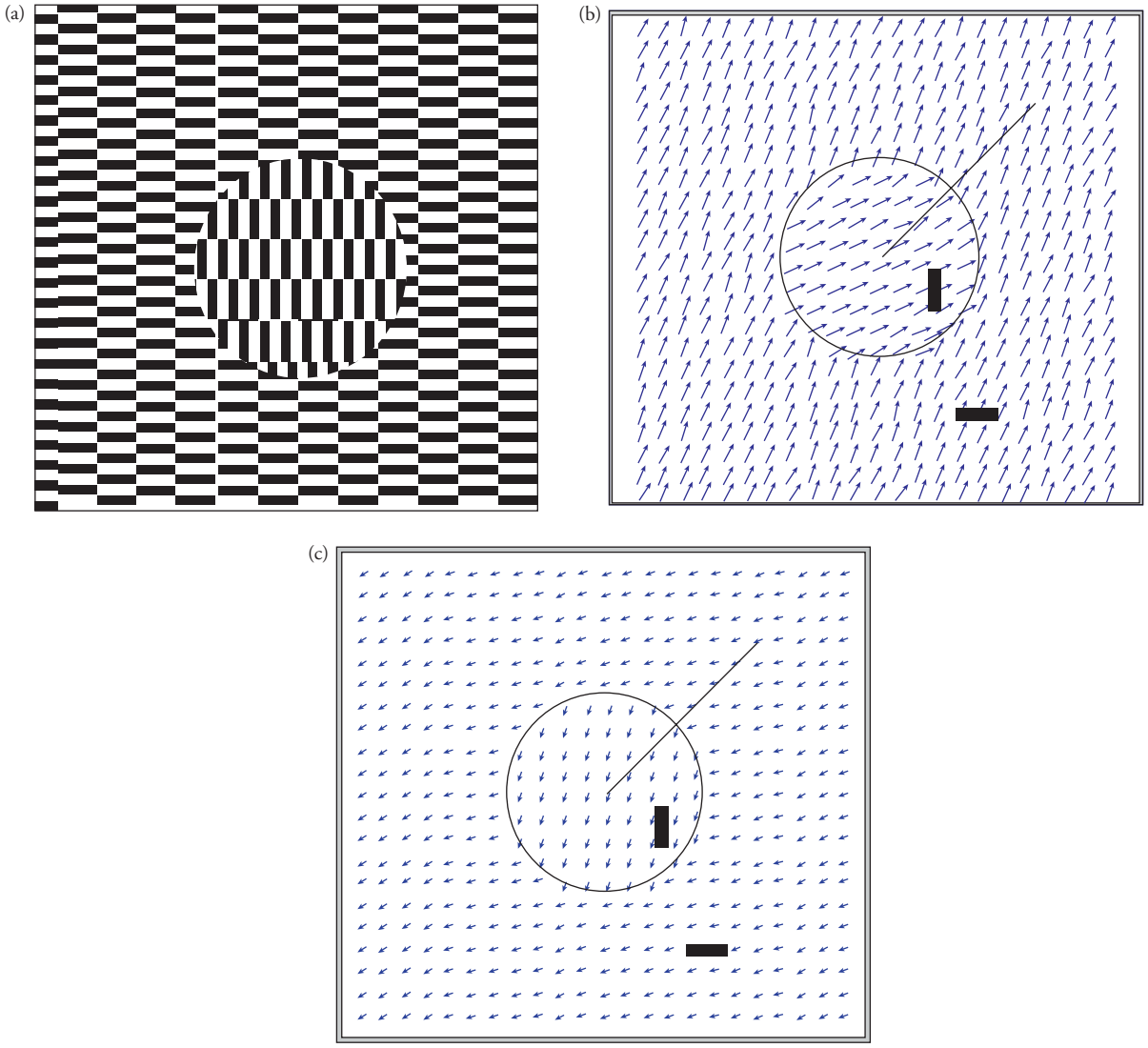


Figure I.6-2. (a) A variant of the Ouch illusion. (b) Estimated optical flow field. (c) The error vector field the difference between the estimated and the veridical motion. The line from the center denotes the direction of the veridical motion. (From Fig. 17: Fermüller, Pless, & Aloimonos, 2000.)

a texture that is dense and thus has small variance. The experiments found that our perception is consistent with this hypothesis.

Finally, we think that the bias not only is a problem of flow estimation but many other visual computations that amount to estimations processes. Early visual processes estimate features, such as lines, points, and image motion (Fermüller & Malm, 2004). All these processes are biased, and as a result the locations of features are perceived erroneously. These effects can account for many classically optical-geometric illusions, as well as the motion illusion explained earlier. The processes of computing the shape of the scene from image features, called *shape from X* computations, are also estimation processes. The bias can account for the erroneous estimation of shape (an underestimation of slant) found in many psychophysical experiments (Ji & Fermüller, 2006) as well as illusions of shape that we created using the insight on the bias (see the Companion Website www.cfar.umd.edu/~fer/optical/).

ERRORS IN NORMAL FLOW ESTIMATION DUE TO CAUSAL FILTERING

Next we look closer at the estimation of normal flow and show that for certain signals and certain image resolutions, one-dimensional image motion is misestimated. The effected signals are asymmetric at some resolution. Take a look at the striking illusion "Rotating Snakes" (Fig. I.6-3a; Kitaoka, 2003). Most observers experience very strong illusory movement when viewing this and similar patterns, such as Figure I.6-3b. The illusory movement is experienced under free viewing conditions when one moves the eyes and is perceived in noncentral vision. For example, while one looks at the upper left of the pattern, motion is perceived in its central and lower right part. From this we see an essential aspect of the effect is that it depends on the image resolution.

The patterns are composed of image patches that have an asymmetric intensity profile. For example, consider

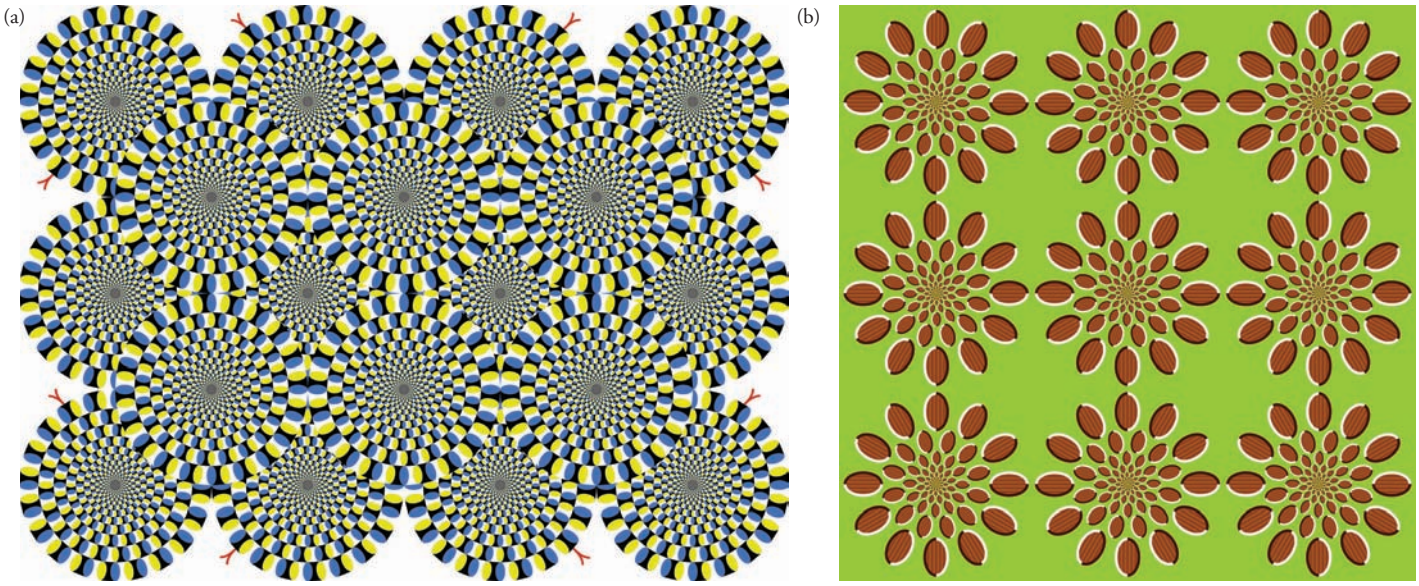


Figure I.6-3. (a) Rotating Snake. (b) Variation of Donguri pattern. In peripheral vision most observer experience rotary movement in both patterns. The direction in the circular arrangements alternates, with counter-clockwise direction in the upper left. (From Kitaoka 2003.)

a narrow slice in the middle region of one of the ovals in Donguri, as shown in Figure I.6-4a. (The Japanese word *donguri* translates to "acorn.") Its monochromatic intensity image can be described as a white and a dark bar (the boundaries of the oval) next to different shades of gray. Referring to Figure I.6-4b, from the highest intensity (the white bar), the intensity drops about twice as much on the right than on the left side. Similarly, from the lowest intensity (the dark bar), the intensity rises about twice as much on the right than on the left. Thus, at the two bars the changes of intensity in the right and left neighborhood are different. Informally, we say that the pattern is asymmetric (at the scale of the size of the bars). Patterns with such intensity profiles create a very strong illusory effect. The perceived movement is a drift from the intensity extremum in the direction of lesser intensity change (i.e., from the white bar to light gray, and from the dark bar to dark gray; Ashida & Kitaoka, 2003).

The analysis is performed using filters as it allows us to analyze the signal under a continuous change in scale. However, a gradient-based approach, observing the changes under smoothing with gradually increasing smoothing and difference kernels, will lead to similar results.

To simplify the analysis, we consider a one-dimensional image with a bar-like structure parallel to the vertical dimension and the motion component perpendicular to the bars. Thus we have a 2D function $I(x,t)$, and the image motion constraint (eq. 2) becomes

$$u\omega_x = -\omega_t, \tag{9}$$

defining a line in the 2D frequency space. The velocity u is obtained from the ratio of the temporal and spatial frequency and amounts to

$$u = -\frac{\omega_t}{\omega_x}. \tag{10}$$

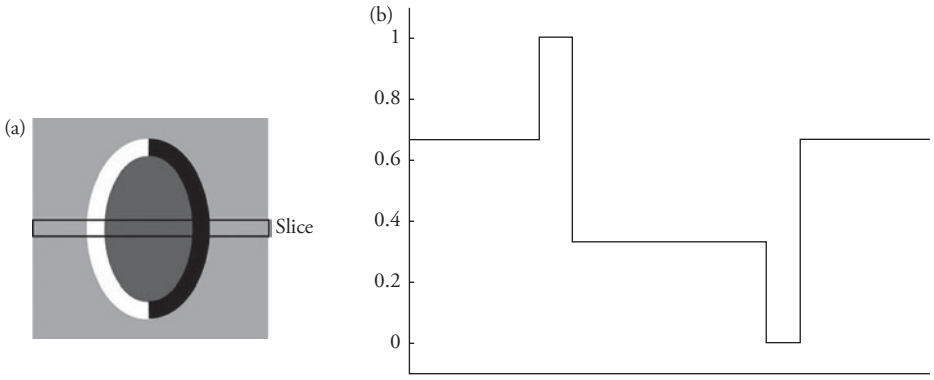


Figure I.6-4. (a) Slice through a patch in the Donguri pattern. (b) Its intensity profile. (From Fig. 3(a)(b) of Fermüller, Ji & Kitaoka, 2010.)

The Filter

The spatiotemporal energy filters for extracting motion are separable in space and time. This just means that the filters can be created as the product between a spatial and a temporal filter. For the analysis this means that the spatiotemporal signal may first be convolved with the spatial filter, and the result may then be convolved with the temporal filter. We follow the common formulation of modeling a filter for detecting the local frequency ω_0, as a complex function

$$g(y) = p(y) \cdot exp(2\pi i \omega_0 y) \tag{11}$$

$exp(2\pi i \omega_0 y)$, called the carrier function, is a complex sinusoid for detecting the signal's component of frequency ω_0, and $p(y)$, called the envelope function, localizes the sinusoid in image space. (Complex functions are needed so we can extract motion independent of the phase of the signal, that is, independent of the position of the signal within the receptive field at certain time and independent of the sign of the contrast [Adelson & Bergen, 1985; Watson & Ahumada, 1985]). Figure I.6-5 illustrates the impulse response function of a biological plausible filter with symmetric envelope in the spatial domain and asymmetric envelope in the time domain (Adelson & Bergen, 1985; Burr & Morrone, 1993). In our model the envelope is a Gaussian in the spatial domain leading to a Gabor filter and a gamma probability density function in the temporal domain (Chen, Wang, & Qian, 2001; Shi, Tsang, & Au, 2004).

The Effect of Filtering on Donguri

Figure I.6-6 illustrates the effect of spatial filtering on Donguri. At frequencies ω_x larger than the width of the bar, the Gabor filter detects the two edges at the left and right of the bar (Fig. I.6-6b). At frequencies significantly smaller than the width of the bar, the Gabor detects the bar (Fig. I.6-6c). The amplitude of the response thus has either one or two well-separated peaks. However, for frequencies of ω_x close to the width of the bar, there is something in between one and two responses. The amplitude function becomes asymmetric with two merging peaks, a larger on the right and a smaller on the left (Fig. I.6-6d). We call these frequencies the *critical frequencies*.

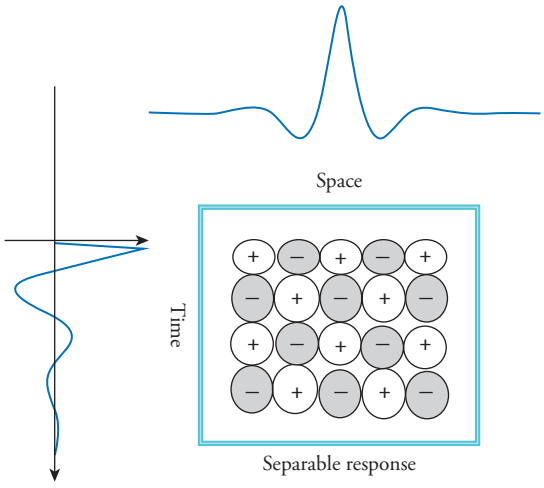


Figure I.6-5. Illustration of biological implementation of spatiotemporal filter. The spatial and temporal impulse responses are shown along the margins. Their product is shown schematically in the center. (Fig. 4, from Fermüller, Ji & Kitaoka, 2010.)

As is well known from the uncertainty principle, there is a limit on the accuracy of localization in image and frequency domain. The Gabor (which is the filter with the best localization in joint image and frequency space) cannot guarantee perfect localization of the signal. The filtered signal should have a local dominant frequency of ω_x, but, because of the "hat" profile of its Gaussian envelope, this will not always be the case. For the critical frequencies the value is quite different from ω_x.

When now estimating on the asymmetric filtered signal, image motion with asymmetric temporal filters, left and right motion are estimated at different values. We can intuitively understand this from the shape of the signals: we convolve the asymmetric spatially filtered signal (Fig. I.6-6d) with larger weight on the right and smaller weight on the left with a temporal filter that in the one case has larger weight on the left and smaller weight on the right and in the other case the filter has larger weight on the right and smaller weight on the left (the mirror reflection). The regions under the two peaks interact with each other during temporal filtering. Referring to Figure I.6-6a, it can be shown through simulation (Fermüller, Ji, & Kitaoka, 2010) that local estimated velocity is overestimated at the right peak for left motion and underestimated at the left peak for right motion. As a result, for the critical frequencies, motion to the left ($u = -1$) leads to larger velocity estimates than motion to the right ($u = 1$).

Explaining the Illusion

The complete explanation of the illusions is as follows. The motion is caused by involuntary fixational eye movements. Work by Murakami, Kitaoka, and Ashida (2006) implicates drift eye movements.[2] The small eye movements cause a change of the image on the retina and trigger the estimation of a motion field. This motion field is due to rigid motion and thus has a certain structure. Under normal conditions the vision system estimates the parameters of the rigid motion and compensates for the image motion field (i.e., the images are stabilized; Murakami, 2004; Murakami & Cavanagh, 1998). Even for asymmetric signals, the vision system can estimate quite accurately the 3D rigid motion using the average of the motion vectors in the patterns. However, as shown earlier, a misestimation occurs at certain locations in the image. Local one-dimensional motion signals are estimated erroneously, and as a result the 2D image motion estimated as shown previously (using eq. 2) is erroneous as well. The difference between the estimated rigid motion field and the erroneously estimated image motion vectors gives rise to residual motion vectors. These residual motion vectors are integrated over time and space, causing the perception of illusory motion in the image.

To evaluate quantitatively the predictive power of the model we performed an experiment with human subjects

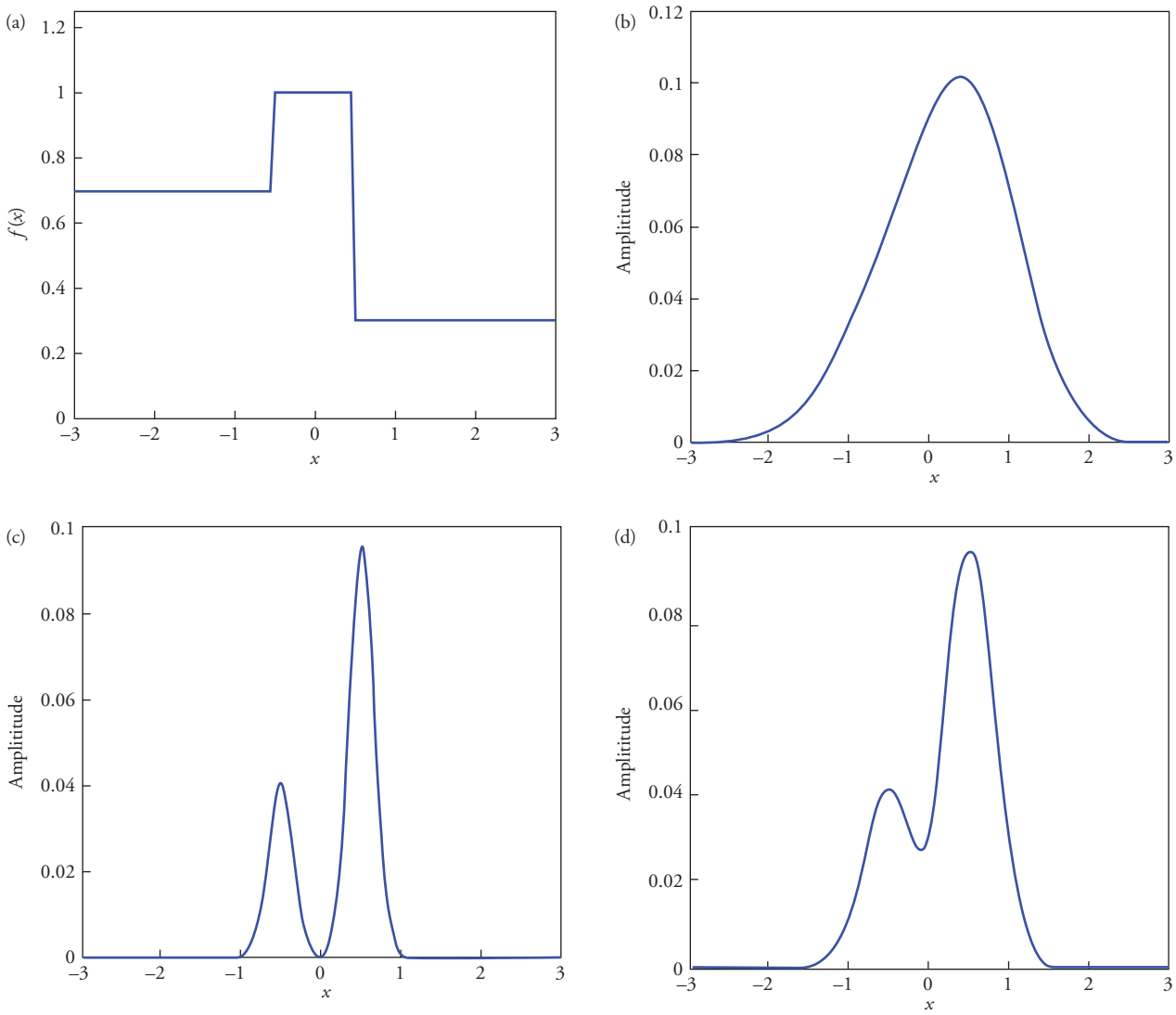


Figure I.6-6. (a) Bar in Donguri pattern. (b–d) Amplitude of bar filtered with Gabor of different frequencies ω_x. (From Fig. 8(a)–(d) of Fermüller, Ji & Kitaoka, 2010.)

(Fermüller et al., 2010). Three different signals of bar-like structures, giving rise to differently strong illusory percepts, were compared in a nulling experiment with opposing real motion. The different signals were arranged on concentric circles shown to naive subjects, who were instructed to adjust the speed of motion in a video until they experienced a stationary pattern. The model was shown to predict well the relative motion speed in the different signals. The model also was shown to account for a series of other illusions, experienced in central and peripheral vision, which Kitaoka and Ashida (2004) call central drift illusions.

Finally, we summarize existing explanations of this illusion, which we broadly classify into two types. The first type, which includes our model and work by Ashida and Kitaoka (2003), suggests that small eye movements cause retinal motion, and the illusion arises due to asymmetric temporal processing. Our model mathematically explains the reason for the erroneous estimation and introduces the concept of the dependence of the estimation at different resolutions of the pattern. The second kind of theory, originating from Faubert and Herbert (1999), suggests that temporal differences in luminance processing produce a motion signal. The theories differ in how this signal is produced. The main idea is that high-contrast regions of the image are transmitted faster than low-contrast regions leading to a motion from high-contrast to low-contrast areas (Conway et al., 2005). Backus and Oruç (2005) suggest that differently strong contrasts and intensities cause different neural response curves over time. Their model introduces the effect of adaptation, which can account for the smooth perception under fixation over a few seconds. This effect may likely exist in addition to the one discussed here.

The illusory perception in the snake pattern is perceived under free viewing conditions and also when flashing the pattern, and many theories seek to explain these effects with a single cause. It is easy to understand that flashing will produce a reverse phi motion (Conway et al., 2005). Reverse phi motion is an image motion effect caused by

reversing the contrast in some frames of a video sequence. However, in our opinion the illusory effect under free viewing and the effect when flickering the pattern are not the same. In support of this, we observed that for the reduced experimental stimuli the latter is experienced much stronger than the former, and it is experienced even by observers who do not experience the effect under free viewing.

CONCLUSIONS

This chapter discussed that, for reasons of statistics and signal processing, it is not possible to estimate correct image motion for all signals. First, temporal image motion filters are causal (i.e., they use data from the past but do not use data from the future). The asymmetry in these filters leads to erroneous estimation of local one-dimensional image motion measurements for asymmetric signals at certain scale. Second, the estimation of 2D image motion using as input one-dimensional motion signals is statistically biased. Because there is noise in the spatiotemporal signal, the estimation cannot be without error. We argue that these limitations in the estimation of flow are integral also to human vision and are the main cause of many optical illusions. We tested our models quantitatively and used them to create new illusory percepts.

Estimation of image motion estimation is the first step in visual motion analysis. The local image motion signals are input to further visual processes. The very basic motion processes include the estimation of our own motion, that is, our head, body motion, and eye movements; the movement of objects in the scene; and the segmentation of the scene into different surfaces. While causal filters and bias in estimation create local erroneous motion signals in the motion illusions discussed here, the strong perception of parts of the image moving with a rotating, sliding, or expanding movement is due to these further processes. But how the interaction between the different processes is implemented is still a topic of research. In our opinion, research into two major questions about optical illusions may help us gain insights.

First, what is the role of motion analysis in the estimation of self-motion and eye-movements? When we move, the image motion on our eyes is due to our body movements, our voluntary and involuntary eye movement, and the movements of objects. The inertial sensors also provide information. But how are the different sources integrated and the different components of movement separated? Under normal circumstances, the images on our eyes are stabilized; we have a stable model of the scene and do not experience micro-movements. As an example of studying an optical illusion, consider the snake illusion. While it can be shown that there are errors on any image motion in asymmetric signals, the illusory perception in the snake illusion has been shown to be due only to the motion signal from drift movements (Kuriki, Ashida, Murakami, & Kitaoka, 2008; Murakami et al., 2006). We speculate that drift movements allow for a better temporal integration of the motion signal when compared to signals from other movements. The drift motion is computed from the local motion signals over the whole visual field. Then the drift is discarded, and the image signals over a time interval are integrated. This is computationally feasible, because the drift motion is mostly a rotation and does not depend on the structure of the scene. By fitting to the whole image motion field a rotational motion field, which depends on only three parameters, local motion vectors can be estimated very accurately and reliably. On the other hand, head motions and scene motions also involve translation, and the image motion field then depends on the scene. Therefore, local motion estimation cannot be that accurate, and integration over a time interval is more difficult.

Second, how do the different processes of early motion perception interact? Insight from computational and biological studies indicate that biological systems implement the estimation of image motion, self-motion, and motion segmentation not in a bottom-up fashion but rather solve them iteratively. As the system moves, it keeps updating its world model. We would argue that first only approximations of image motion and rough segmentation are derived, and self-motion is computed. Then estimates of 3D motion and segmentation help with estimating better image motion and in turn more accurate segmentation at the next time instance. Furthermore, at the level of segmentation, image motion is combined with static cues such as contour, color, and texture to obtain the boundaries between surfaces. As an example of using an illusion to gain insight, consider the Leviant illusion (Zeki, Watson, & Frackowiak, 1993). We hypothesize that the main reason for this illusion lies in the interaction between the three processes of image motion, ego motion, and segmentation, and in Fermüller, Pless, and Aloimonos (1997), we showed that we can create variants of this illusion from patterns that have a special role in ego-motion estimation.

NOTES

1. A third class of methods used in the computational literature, models optical flow as correlation of regions followed by interpolation; this does not proceed in the same two steps and requires a somewhat different modeling (Fermüller et al., 2001).
2. The drift movements, one of the three fixational eye movements, are defined as incessant random fluctuations at about 1 to 30 Hz, quite large (~10 min of visual angle) and fast (up to $2 - 3^{\circ}/s$; Eizenman, Hallett, & Frecker, 1985). They are more rigorous after a saccade (Ross, Morrone, Goldberg, & Burr, 2001) than during steady fixation.

REFERENCES

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, *2*, 284–299.

Adelson, E. H., & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, *300*, 523–525.

Anstis, S. (1970). Phi movement as a subtraction process. *Vision Research*, *10*, 1411–1430.

Anstis, S. (2003). Moving objects appear to slow down at low contrasts. *Neural Networks 16*, 933–938.

Anstis, S., & Rogers, B. (1986). Illusory continuous motion from oscillating positive-negative patterns: Implications for motion perception. *Perception*, *15*, 627–640.

Ashida, H., & Kitaoka, A. (2003). A gradient-based model of the peripheral drift illusion. *Perception*, *32*(supplement), 106.

Backus, B. T., & Oruc, I. (2005). Illusory motion from change over time in the response to contrast and luminance. *Journal of Vision*, *5*(11), 1055–1069.

Burr, D., & Morrone, M. (1993). Impulse-response functions for chromatic and achromatic stimuli. *Journal of the Optical Society of America A*, *10*, 1706–1713.

Chen, Y., Wang, Y., & Qian, N. (2001). Modeling V1 disparity tuning to time-varying stimuli. *Journal of Neurophysiology*, *86*, 143–155.

Chubb, C., & Sperling, G. (1988). Drift balanced random dot stimuli: A general basis for studying non-Fourier motion. *Journal of the Optical Society of America A*, *5*, 1986–2007.

Conway, B., Kitaoka, A., Yazdanbakhsh, A., Pack, C., & Livingstone, M. (2005). Neural basis for a powerful static motion illusion. *Journal of Neuroscience*, *25*, 5651–5656.

Eizenman, M., Hallett, P., & Frecker, R. (1985). Power spectra for ocular drift and tremor. *Vision Research*, *25*, 1635–1640.

Exner, S. (1875). Experimentelle Untersuchung der einfachsten psychischen Processe. *Pflügers Archiv für die gesamte Physiologie des Menschen und der Tiere*, *1*, 403–472.

Faubert, J., & Herbert, A. M. (1999). The peripheral drift illusion: A motion illusion in the visual periphery. *Perception*, *28*, 617–621.

Fermüller, C., Ji, H., & Kitaoka, A. (2010). Illusory motion due to causal time filtering. *Vision Research*, *50*(3), 315–329.

Fermüller, C., & Malm, H. (2004). Uncertainty in visual processes predicts geometrical optical illusions. *Vision Research*, *44*, 727–749.

Fermüller, C., Pless, R., & Aloimonos, Y. (1997). Families of stationary patterns producing illusory movement: Insights into the visual system. *Proceedings of the Royal Society of London*, *B*, *264*, 795–806.

Fermüller, C., Pless, R., & Aloimonos, Y. (2000). The Ouchi illusion as an artifact of biased flow estimation. *Vision Research*, *40*, 77–96.

Fermüller, C., Shulman, D., & Aloimonos, Y. (2001). The statistics of optical flow. *Computer Vision and Image Understanding*, *82*(1), 1–32.

Fuller, W. (1997). Estimated true values for errors-in-variables models. In S. van Huffel (Ed.), *Recent advances in total least squares techniques and errors-in-variables modeling*. Philadelphia: Society for Industrial and Applied Mathematics.

Gillam, B. (1998). Illusions at century's end. In J. Hochberg (Ed.), *Perception and cognition at century's end* (pp. 95–136). New York: Academic Press.

Gregory, R. L. (1968). Perceptual illusions and brain models. *Proc Royal Sci London: Series B − Biological Sciences*, *171*, 279–296.

Grossberg, S., & Todorovic, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena. *Perception & Psychophysics*, *43*, 241–277.

Helmholtz, H. L. F. von. (1962). *Treatise on physiological optics*. Vol. III. (J. P. C Southall, Trans.). New York: Dover.

Horn, B. K. P., & Schunck, B. (1981). Determining optical flow. *Artificial Intelligence*, *17*, 185–203.

Ji, H., & Fermüller, C. (2006). Noise causes slant underestimation in stereo and motion. *Vision Research*, *46*(19), 3105–3120.

Khang, B.-G., & Essock, E. A. (1997). Apparent relative motion from a checkerboard surround. *Perception*, *26*, 831–846.

Kitaoka, A. (2003). Akiyoshi's illusion pages. http://www.ritsumei.ac.jp/~akitaoka/index-e.html

Kitaoka, A., & Ashida, H. (2004, January). A new anomalous motion illusion: The "central drift illusion." Paper presented at the winter meeting of the Vision Society of Japan. Kogakuin University, Shinjuku, Tokyo.

Kuriki, I., Ashida, H., Murakami, I., & Kitaoka, A. (2008). Functional brain imaging of the rotating snakes illusion by fMRI. *Journal of Vision*, *8*(10), 1–10.

Laughlin, S. (1981). Neural principles in the peripheral visual systems of invertebrates. In H. Autrum (Ed.), *Handbook of sensory physiology* (Vol. 7, pp. 133–280). New York: Springer.

Marr, D. (1982). *Vision*. San Francisco, CA: W. H. Freeman.

Marr, D., & Ullman, S. (1981). Directional selectivity and its use in early visual processing. *Proceedings of the Royal Society of London*, *B*, *211*, 151–180.

Mingolla, E. (2003). Neural models of motion integration and segmentation. *Neural Networks*, *16*(5–6), 939–945.

Morrone, M., & Burr, D. (1988). Feature detection in human vision: A phase dependent energy model. *Proceedings of the Royal Society of London*, *B*, *225*, 221–245.

Murakami, I. (2004). Correlations between fixation stability and visual motion sensitivity. *Vision Research*, *44*, 251–261.

Murakami, I., & Cavanagh, P. (1998). A jitter after-effect reveals motion-based stabilization of vision. *Nature*, *395*, 798–801.

Murakami, I., Kitaoka, A., & Ashida, H. (2006). A positive correlation between fixation instability and the strength of illusory motion in a static display. *Vision Research*, *46*, 2421–2431.

Neumann, H., & Mingolla, E. (2003). Contour and surface perception. In M. Arbib (Ed.), *Handbook of brain theory and neural networks* (Vol. II, pp. 217–276). Cambridge, MA: MIT Press.

Palmer, S. E. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA: MIT Press.

Pinna, B., & Brelstaff, G. J. (2000). A new visual illusion of relative motion. *Vision Research*, *40*(16), 2091–2096.

Ross, J., Morrone, M., Goldberg, M., & Burr, D. (2001). Changes in visual perception at the time of saccades. *Trends in Neurosciences*, *24*(2), 113–121.

Shi, B. E., Tsang, E. K. C., & Au, P. S. P. (2004). An on-off temporal filter circuit for visual motion analysis. In *Circuits and Systems, 2004: ISCAS '04. Proceedings of the 2004 International Symposium* (Vol. 3, pp. 85–88). Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Spillmann, L., Heitger, F., & Schuller, S. (1986). Apparent displacement and phase unlocking in checkerboard patterns. Paper presented at the 9th European Conference on Visual Perception. Bad Nauheim, Germany.

Spillmann, L., Tulunay-Keesey, U., & Olson, J. (1993). Apparent floating motion in normal and stabilized vision. *Investigative Ophthalmology and Visual Science*, *34*(4), 1031.

Srinivasan, M., Laughlin, S., & Dubs, A. (1982). Predictive coding: A fresh view of inhibition in the retina. *Proceedings of the Royal Society of London*, *B*, *216*, 427–459.

van Huffel, S., & Vandewalle, J. (1991). *The total least squares problem: Computational aspects and analysis*. Philadelphia: Society for Industrial and Applied Mathematics.

Wallach, H. (1935). Über visuell wahrgenommene Bewegungsrichtung. *Psychologische Forschung*, *20*, 325–380.

Watson, A. B., & Ahumada, A. J. (1985). Model of human visual motion sensing. *Journal of the Optical Society of America*, *2*, 322–342.

Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung. *Zeitschrift für Psychologie*, *61*, 161–265. (English translation in T. Shipley [Ed.], *Classics in Psychology*, New York: Philosophical Library, 1961)

Yo, C., & Wilson, H. R. (1992). Moving 2D patterns capture the perceived direction of both lower and higher spatial frequencies. *Vision Research*, *32*, 1263–1270.

Zeki, S. M., Watson, J. D. G., & Frackowiak, R. S. J. (1993). Going beyond the information given: The relation of illusory visual motion to brain activity. *Proceedings of the Royal Society of London*, *B*, *252*, 215–222.

Chapter 7
The Visual World as Illusion
The Ones We Know and the Ones We Don't

Stephen Grossberg

EXPECTATION, IMAGINATION, AND ILLUSION

When you open your eyes in the morning, you usually see what you *expect* to see. Often it will be your bedroom, with things where you left them before you went to sleep. What if you opened your eyes and unexpectedly found yourself in a steaming tropical jungle? Why do we have *expectations* about what is about to happen to us? Why do we get *surprised* when something unexpected happens to us? More generally, why are we Intentional Beings who are always projecting our expectations into the future? How does having such expectations help us to fantasize and plan events that have not yet occurred? Without this ability, all creative thought would be impossible, and we could not imagine different possible futures for ourselves or our hopes and fears for them.

What is the difference between having a fantasy and experiencing what is really there? What is the difference between illusion and reality? What goes wrong when we lose control over our fantasies and hallucinate objects and events that are not really there? Given that vivid hallucinations are possible, especially in mental disorders like schizophrenia, how can we ever be sure that an experience is really happening and is not just a particularly vivid hallucination? If there a fundamental difference between reality, fantasy, and illusion, then what is it?

SEEING, RECOGNIZING, AND CONSCIOUSNESS

Once we mention illusion, it is hard not to talk about consciousness. Our consciousness of visual events is particularly vivid to us because we get so much information about the world from our ability to see. But why do we bother to see? One attractive hypothesis is that we see objects in order to recognize them, so that we can then act upon them to test our expectations or fulfill our desires. But do we really need to *see* something in order to *recognize* it? The answer is no. For example, when we look at the image in Figure I.7-1a, we can all recognize that there is a vertical line, or boundary, of some sort that is interpolated between the array of horizontal black lines. But we do not *see* this line, because it does not have a different brightness or color from its background. Nor does it separate the image into regions of different brightness or color. We *recognize* that the line is there, but it is invisible, or not seen. Such a percept is called *amodal* in this chapter. A *modal* percept is one that does carry some visible brightness or color difference. This use of the term *amodal* generalizes the more traditional use by authors such as Albert Michotte, Gabio Metelli, and Gaetano Kanizsa because the mechanistic analysis of how boundaries and surfaces interact, which is summarized later in this chapter, supports a more general terminology in which perceptual boundaries are formed without any corresponding visible surface qualia.

Simple images like this turn our naive views about the function of seeing upside down. For example, one introspectively appealing answer to the question "Why do we see?" is that "We see things in order to recognize them." But we can recognize the vertical boundary of an offset grating without seeing it. The offset grating is a counterexample to the hypothesis that seeing is needed for recognizing. But if we do not always need to see things in order to recognize them, then why do we bother to see at all? Once we start making observations like this, we realize that things are not what they seem and that we need a different way of thinking in order to understand how we consciously see the world.

ILLUSION AND REALITY

The percept that we experience when viewing Figure I.7-1 shows that we can recognize, and be conscious of, events that are invisible. Many people, including distinguished philosophers, have written about visual consciousness as if it includes only consciousness-of-qualia, or of visible sensations. Figure I.7-1 shows that this is a wrong, or at least overly simple, idea.

This percept of an invisible vertical boundary is also of interest because it is a *visual illusion*—it is not "really there" in the image. Why can we see or recognize illusions? How much of what we see is an illusion? Many Eastern thinkers, such as in Buddhism, have spoken about the world-as-illusion. How correct are they? Do they think we see illusions for the right reasons? Moreover, if our consciousness can include even such bizarre experiences as invisible illusions, then why do we not we get hopelessly confused between illusion and reality all the time? Is there some adaptive property of our brains that implies that we can be conscious of

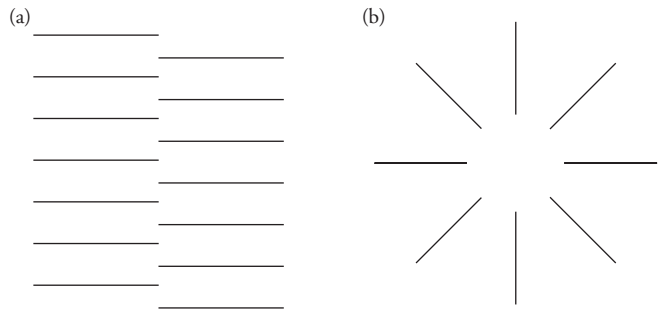


Figure I.7-1. (a) Offset grating. The vertical boundary that emerges between the horizontal black edges can be recognized but not seen. (b) The Ehrenstein illusion. The ends of the black lines induce an illusory circular boundary as well as "brightness buttons" that fill in within the circle to create a percept of enhanced brightness relative to the level of filled-in brightness exterior to the circle.

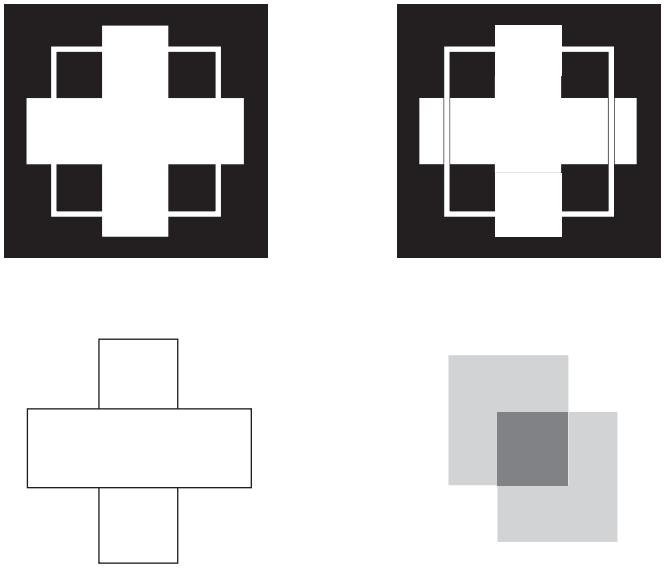


Figure I.7-2. Kanizsa stratification figures are in the top row. The left figure usually looks like a white cross in front of a partially occluded white square, but can switch to look like the square in front of a partially occluded cross. The white region that is shared by the square and the cross is "captured" by the surface form that appears to be in front. In the right figure, complete vertical lines of the square in front of the cross leads to an unlikely percept of a cross that interleaves a square in depth. Bayesian theorists claim that what we see is likelihood statistics. However, this figure causes the highly unlikely percept of a cross both in front of and behind the square. The left image in the second row contains three abutting rectangles but is perceived as a horizontal bar that partially occludes a vertical bar just behind it in depth. The occluded vertical surface is invisible, or amodal, but can nonetheless be recognized. The right image in the second row is also composed of three regions. It generates a bistable percept of a transparent square in front of a background square. (Top row, left column figure is from "Nuove Ricerche Sperimentali Sulla Totalizzazione Percettiva," by G. Petter, 1956, Rivista di Psicologia. 50. Copyright 1956 by Giunti: Gruppo Editoriale. Adapted with permission. Top row, right column figure is from "Seeing and Thinking," by G. Kanizsa, 1985, Acta Psycologia, 59. Copyright 1985 by Elsevier Science. Adapted with permission.)

invisible illusions, without preventing us from experiencing "reality" most of the time? If so, what is it? In fact, in all cases that have been analyzed by our neural theory of how the brain sees, visual illusions arise from adaptive properties of our brains that are needed for survival. More generally, many of the percepts that we believe to be "real" and that we use to survive, are, strictly speaking, visual illusions that are actively constructed from the visual information that we receive from the world through our retinas.

ART AND MOVIES ARE SEEN USING BOUNDARIES AND SURFACES

Before explaining this, let me emphasize that the visual arts and the entertainment industry depend greatly upon visual illusions. Whenever we look at a photograph, painting, movie screen, or computer monitor, we respond to a two-dimensional (2D) picture by constructing a three-dimensional (3D) percept—that is, the picture is experienced as a scene in depth. The human urge to represent the world using 2D pictures dates back at least to Paleolithic times. Artists from ancient to modern times have sought to master how a few lines or color patches on a flat surface can induce mental representations of multiple objects in a scene, notably of occluding objects in front of the objects that they partially occlude. Neural models of seeing have clarified how our ability to experience the world through a picture, or series of pictures in a movie, derives from more basic properties of how we see the world in depth (Grossberg, 1994, 1997).

Figure I.7-2 illustrates the claim that percepts derived from pictures are often illusions. The lower left image in Figure I.7-2 shows three rectangular shapes that abut one another. Our percept of this image irresistibly suggests a different interpretation, however. We perceive a horizontal bar lying in front of a partially occluded vertical bar that is amodally completed behind it. The alternative percept of two rectangles abutting a horizontal bar is much less frequent, although it provides a more literal description of the image.

Even this simple image raises deep questions: Why do the horizontal edges, or boundaries, that are shared by two abutting rectangles seem to "belong" to the horizontal rectangle? When these shared horizontal boundaries are attributed to the horizontal bar, the vertical boundaries of the two rectangles are somehow freed from interference by these horizontal boundaries. The vertical boundaries can then cooperate to induce amodal vertical boundaries "behind" the horizontal bar. These amodal boundaries connect the vertical boundaries of the two rectangles to form a percept of a partially occluded vertical bar. But how can a 2D picture generate a percept in which anything is *behind* anything else? How do two dimensions create a percept in three dimensions? Moreover, how do these perceptual properties relate to classical mathematical concepts of geometry? Whatever its origins, this sort of phenomenon is clearly essential for our ability to recognize the world through pictures.

The two images in the top row of Figure I.7-2 provide another example of how our brains can interpret 2D pictures as events in three dimensions. These images were introduced by the great Italian psychologist Gaetano Kanizsa, who was one of the most insightful contributors to visual perception until his death in 1993. They are called

stratification images because they generate a percept of an occluding object in front of an occluded object. Thus, although the left image is just a white region on a black background, it generates a percept either of a white cross in front of a white outline square, or of a white outline square in front of a white cross. The former percept usually occurs, but the percept can intermittently switch between these two interpretations. When a percept switches like this between two different interpretations, it is said to be *bistable*.

Stratification percepts raise many questions: How does the boundary of the cross get completed across the white regions where it intersects the square? How does the white color within the completed cross boundary get "captured" into a cross-shaped surface that is enclosed by this boundary? How is this cross assigned to a depth closer than that of the square? How does the rest of the white region get assigned to the square at this more distant depth? How do the incomplete square boundaries get amodally completed "behind" the cross to facilitate recognition of the square? Why do not the completed square boundaries capture the white color within them where they intersect the cross? When the percept switches, so that the square appears in front of the cross, why do the completed square boundaries succeed in capturing the white color that previously was attributed to the cross, whereas the completed cross boundaries are merely amodally completed behind the square? Why is the percept bistable? And why does the cross win over the square more often than not?

One might at first think that such percepts arise because of prior learning about familiar objects in the world, like crosses and squares, as the large psychophysical literature on perceptual learning illustrates. Although visual experiences do influence how the brain develops and learns, something different is going on when we perceive images such as those in Figure I.7-2. For example, in the upper right image of Figure I.7-2, two pairs of vertical black lines create boundaries that help to capture the intervening white color into a vertical surface of the square. This vertical region seems to lie in front of the horizontal bars of the cross. This effect influences the percept only locally, however. When we look to the middle of the picture, we see that the vertical bars of the cross are still completed in front of the square. The vertical bars of the cross appear to lie in front of the cross, even though the horizontal bars of the cross appear to lie in back of the cross. Kanizsa used this image to argue that knowledge about the world could not determine such percepts because, surely, our many experiences with crosses and squares would, with high probability, place the cross behind the square *everywhere* once it is forced to be behind the vertical bars of the square. This is not, however, what we perceive. We perceive, instead, the highly unlikely situation in which the cross appears to be deformed to lie both in front and in back of the square. The famous graphic artist M. C. Escher used such unlikely properties to create fascinating images that look sensible locally but are globally impossible.

These 2D images in Figure I.7-2 all generate 3D percepts, and the upper left image shows that such percepts can be bistable, with different 3D perceptual interpretations of the same picture alternating through time. The thickness of the limbs of the cross relative to those of the square are responsible for the cross boundaries appearing in front of the square boundaries most of the time. In response to the lower right image in Figure I.7-2, again, two 3D percepts can alternate through time. In this case, the nearer square looks transparent because one can see the farther square behind it. This percept is due to the relative luminances of the various parts of the image, since changing these luminances can create either a 2D opaque percept or a 3D transparent percept that is not bistable. Thus, either geometrical properties of an image, as in the upper left image, or contrastive properties of the image, as in the lower right image, can cause 3D bistable percepts to be seen.

All of these percepts, and many more like them, have been explained and simulated using the form-and-color-and-depth (FACADE) theory of how the visual cortex accomplishes 3D vision and figure–ground perception, and its refinement as the 3D LAMINART theory of how the laminar circuits of visual cortex carry out these perceptual functions (Cao & Grossberg, 2005, 2012; Fang & Grossberg, 2009; Grossberg, 1987, 1994, 1997; Grossberg & Howe, 2003; Grossberg & McLoughlin, 1997; Grossberg & Pessoa, 1998; Grossberg & Swaminathan, 2004; Grossberg & Yazdanbakhsh, 2005; Grossberg, Yazadanbakhsh, Cao, & Swaminathan, 2008; Kelly & Grossberg, 2000). In particular, Grossberg (1997) predicted how interactions between figural geometry and contrast can change what we see, Kelly and Grossberg simulated various forms of these percepts, and Grossberg and Yazdanbakhsh explained and simulated how both opaque and transparent percepts can occur in response to different combinations of contrasts in the lower right figure.

These figures illustrate a controversy about how we see and recognize the world that has continued to the present time. The alternative view—that we see what we expect to see because of prior learning—also has a lot of evidence to support it. Hermann von Helmholtz, one of the greatest scientists of the 19th century, proposed that we see using *unconscious inferences*, or learned expectations. Helmholtz was a precursor of the use of Bayesian statistics by some vision scientists today. In contrast, Kanizsa provided hundreds of brilliant counterexamples to Helmholtz's hypothesis using images, such as the ones in Figure I.7-3 that violate expectations. Kanizsa emphasized the power of *bottom-up* visual processes, such as filtering, grouping, and filling-in that act directly on visual scenes and images at early stages of brain processing, whereas Helmholtz called attention to the impact of *top-down* processes, such as learned expectations, that are activated by processes deeper in the brain and are matched against these more peripheral processes.

Helmholtz and Kaniza emphasized important, but complementary, types of visual processes. Visual illusions can be created by both bottom-up and top-down processes. A more comprehensive understanding has arisen from neural models that show how bottom-up and top-down processes work together to generate an attentive consensus, or resonance, between bottom-up signals that are derived from the world and what we expect to see. How such a consensus arises also helps to explain how we can rapidly learn about the world throughout life and become conscious of events in it without being forced to unselectively forget already learned knowledge as we try to learn

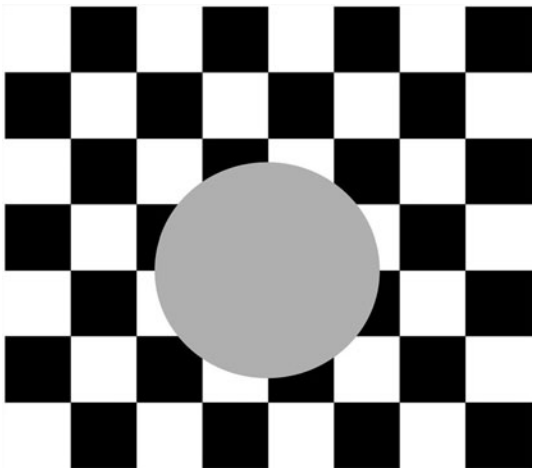


Figure I.7-3. Two gray disks are superimposed on a black-and-white checkerboard grating. The gray disk to the upper left seems to partially occlude a white cross, whereas the disk to the lower right seems to occlude a black cross. Due to the regularity of the checkerboard and our prior experiences with checkerboards, a Bayesian account would have predicted that a black square is occluded at the upper left and a white square at the lower right, but this is not what is perceived.

new things (Cao, Grossberg, & Markowitz, 2011; Carpenter & Grossberg, 1991; Fazl, Grossberg, & Mingolla, 2009; Grossberg, 1980, 2007; Grossberg, Markowitz, & Cao, 2011; Raizada & Grossberg, 2003). Helmholtz and Kanizsa could not fully make these connections because they did not have some of the critical intuitive concepts and none of the mathematics needed to express them clearly.

WHY DO WE BOTHER TO SEE?

Earlier we introduced the distinction between seeing and recognizing, which is classically sometimes described as the distinction between *seeing* and *knowing*. Figure I.7-1a shows that we can consciously recognize the offset grating even though we cannot see it. The image in Figure I.7-1b consists of a series of concentric black lines that face the same central point on the paper. Our percept of this figure is one in which we both see *and* recognize a circle that abuts the ends of the black lines. We *see* this circle because it separates the brighter white disk within the circle from the darker background that lies outside it. This circle is a visual illusion, since neither the circle nor the bright disk within it is part of the picture itself. It is called the Ehrenstein illusion in honor of the German perceptual scientist Walter Ehrenstein, who studied it.

We next propose answers to the following questions:

- Why are there so many visual illusions?
- How does illusion differ from reality?
- Why do illusory percepts not cause us a lot of trouble in successfully adapting to the "real" world?

Illusions provide vital clues about how the brain successfully adapts to the real world, and many percepts and other experiences that we believe to be real are, in actuality, mentally constructed illusions that just happen to match the structure of the world. How this happens clarifies why the debate about the difference between illusion and reality has gone on for so long, why many cultures have emphasized mystical states of mind that exploit the constructive properties of the mind, and why the long-held proposal of artificial intelligence that brains may be modeled by the architecture of a Von Neumann computer is wrong. The next section suggests that our percept of the offset grating in Figure I.7-1a reflects a general fact about how we perceive the "real" world.

ALL BOUNDARIES ARE INVISIBLE

The percept of a vertical boundary in the offset grating shows us that "some boundaries are invisible." In fact, "*all* boundaries are invisible," in the sense that the boundary system does not directly generate a percept of visible qualia such as brightnesses and colors (Cohen & Grossberg, 1984; Grossberg, 1984; Grossberg & Mingolla, 1985a, 1985b). The following discussion explains what perceptual boundaries are, how the brain computes them, and how boundaries reflect a general principle of brain design.

First, consider Figure I.7-3. This famous image was introduced by Kanizsa to argue against the position that all seeing is based on knowledge-based hypothesis testing. In particular, the two gray disks are perceived to partially occlude a white cross and a black cross, rather than the checkerboard squares that would be expected from experience. Figure I.7-3 can also be used to understand a different issue. Let your eye wander along the circular boundary of one of the two gray disks. At successive moments, the boundary may abut a white square, then a black one, then a white one, and so on. Where the square is white, the *contrast* between the background and the disk goes from white to gray, that is, from a lighter to a darker shade of gray. Where the square is black, the contrast goes from black to gray, or from a darker to a lighter shade of gray. As your eye traverses the boundary, the contrast keeps flipping between light to dark and dark to light. Despite these contrast reversals, we perceive a single continuous boundary surrounding the gray disk. The boundary is not broken into two sets of four pieces, depending on whether the direction of contrast goes from light to dark or from dark to light. This would have been the result if all that happened in perceiving the circular boundary was for cells that are sensitive to a light-to-dark contrast, or to a dark-to-light contrast, but not both, responded along the circle. Instead, this percept illustrates that the brain pools, or adds, signals from cells that are sensitive to both of these contrast polarities, at every position, in order to detect the boundaries of objects, irrespective of their direction of contrast. The same thing may be said about the checkerboard background of the image: We can recognize long vertical boundaries from the top to the bottom of the image, even though these boundaries are formed from alternating contrast polarities. We can also recognize the long horizontal boundaries of the checkerboard in the same way.

The punch line is: if the brain adds together signals from opposite contrast polarities in order to build boundaries around objects, then these boundaries cannot tell the difference between light and dark, and thus cannot carry a

visible signal of lightness or darkness. In other words, *all boundaries are invisible* because the brain needs to build boundaries around objects that lie in front of textured backgrounds whose relative contrast with respect to the object can reverse as the boundary is traversed.

Apart from having serious problems with object recognition if boundaries did not pool opposite contrasts, there would also be serious perceptual consequences. For example, if a connected boundary did not form around objects, then its lightnesses and colors could spread, through a surface filling-in process, between objects in a scene and we could not properly see object surfaces.

This contrast-pooling property occurs at an early stage in visual cortical processing, namely, at the complex cells in layer 2/3 of cortical area V1. Perhaps the simplest example of this boundary-pooling property can be perceived by viewing Figure I.7-4. The image to the left of Figure I.7-4 is perhaps the most famous image due to Kanizsa. This image consists of four black pac-man, or pie-shaped, figures on white paper. The percept of the square that abuts the pac-men interiors is a visual illusion that is called the Kanizsa square. The enhanced brightness of the square is also an illusion.

Now consider the image to the right of Figure I.7-4. The percept that it elicits is called a "reverse-contrast Kanizsa square." Again a square boundary can be *recognized*. If the contrasts of the gray background relative to the white and black pac-men are chosen correctly, then we will *not see* a brighter or darker square within the square boundary. In this percept, the two vertical boundaries of the square are formed between pac-man edges of opposite contrast polarity: white to gray (light to dark) and black to gray (dark to light). To do this, the brain pools signals from opposite directions of contrast to form an object boundary, albeit the boundary of an illusory object. In addition, the two horizontal boundaries of the square are formed between like directions of contrast, and both sets of boundaries join together to form the square. Thus both like-contrast and opposite-contrast polarities are pooled together to form an object's boundary, thereby illustrating once again why all boundaries are invisible. In other words, boundaries are *insensitive to contrast polarity*, in the sense that they pool signals from cells that are sensitive to opposite contrast polarities.

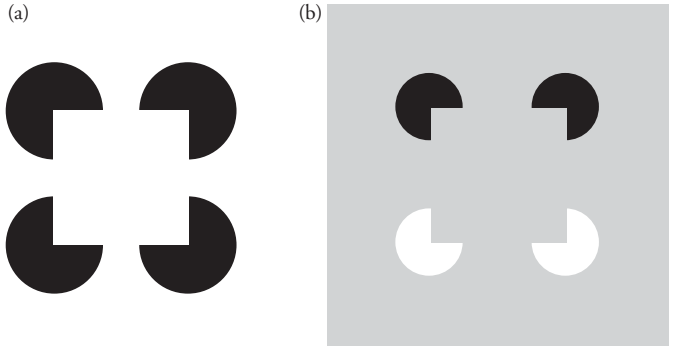


Figure I.7-4. A Kanizsa square (a) and reverse-contrast Kanizsa square (b) illustrate complementary properties of perceptual boundaries and surfaces. Whereas the Kanizsa square may appear brighter than its background, the reverse-contrast Kanizsa square may be recognized, but not seen, if the interior and exterior filled-in gray brightnesses are the same due to the balance between black-to-gray and white-to-gray pac man inducers of filling-in.

If, in fact, all boundaries are invisible, then how do we see the Kanizsa square? How, indeed, do we see anything? Something else must also be going on that lets us see the visible world of surface brightnesses and colors. What is this additional surface perception process, and how is it related to the process whereby the brain forms boundaries? We next build up an answer to these basic questions in several simple stages.

LIVING WITH YOUR BLIND SPOT

To get started, let us first acknowledge that perceptual boundaries and surfaces are not just high-level constructs that have to do with sophisticated processes like object recognition or hypothesis testing. Boundaries and surfaces play a role in processing signals from the very earliest stage of visual processing, the eye itself.

The top of Figure I.7-5 shows a cross-sectional view through an eye. Light streams in from the left, passes through the lens and vitreous humor, and finally hits the photosensitive retina. A ray of light is shown hitting the most sensitive regions of the retina, which is called the fovea. The fovea is the region of the retina with the highest acuity and the highest concentration of color-sensitive photoreceptors. The fovea processes a region of visual space that is approximately 1° wide. Retinal acuity gradually decreases with distance from the fovea. Our eyes move in our heads to direct our sensitive foveas to new points of interest.

Although the retina is a miracle of biological design, it is also a very *noisy* detector. Without the help of additional processing by the brain, retinal noise could easily defeat our ability to see. The bottom of Figure I.7-5 illustrates one source of noise. This figure looks down on the surface of the retina. To the left of the figure is the fovea. To the right, is the *blind spot*. The blind spot is a region of the retina that is totally blind; it has no light-sensitive photodetectors. The blind spot exists because of the way in which the photodetectors in other parts of the retina send light-activated signals to the brain along nerve pathways, called axons. These axons are brought together in the optic nerve, as shown at the top of Figure I.7-5. The blind spot lies right in front of the place where all the axons come together to be bundled into the optic nerve.

The blind spot is large: notice how big it is relative to the fovea. Despite this fact, many of us manage to go through life without ever realizing that we have a large blind hole in our retinas. We certainly do not see a roving hole in the world as our eyes scan a scene. Why not? In addition to the blind spot, Figure I.7-5 shows that a lot of veins lie between the source of light and the retina. These veins nourish the eye, but we do not see these veins either.

There is a second sense in which retinal processing is noisy. When light hits the retinal photodetectors, it causes an electrical response in them. This response is processed by several layers of retinal cells before the processed signals are collected in the optic nerve and sent to the brain. However, these retinal layers are upside-down: light passes through all of them before it ever hits the photodetectors and activates them. All these layers of cells can absorb and scatter the light before it ever reaches the photodetectors.

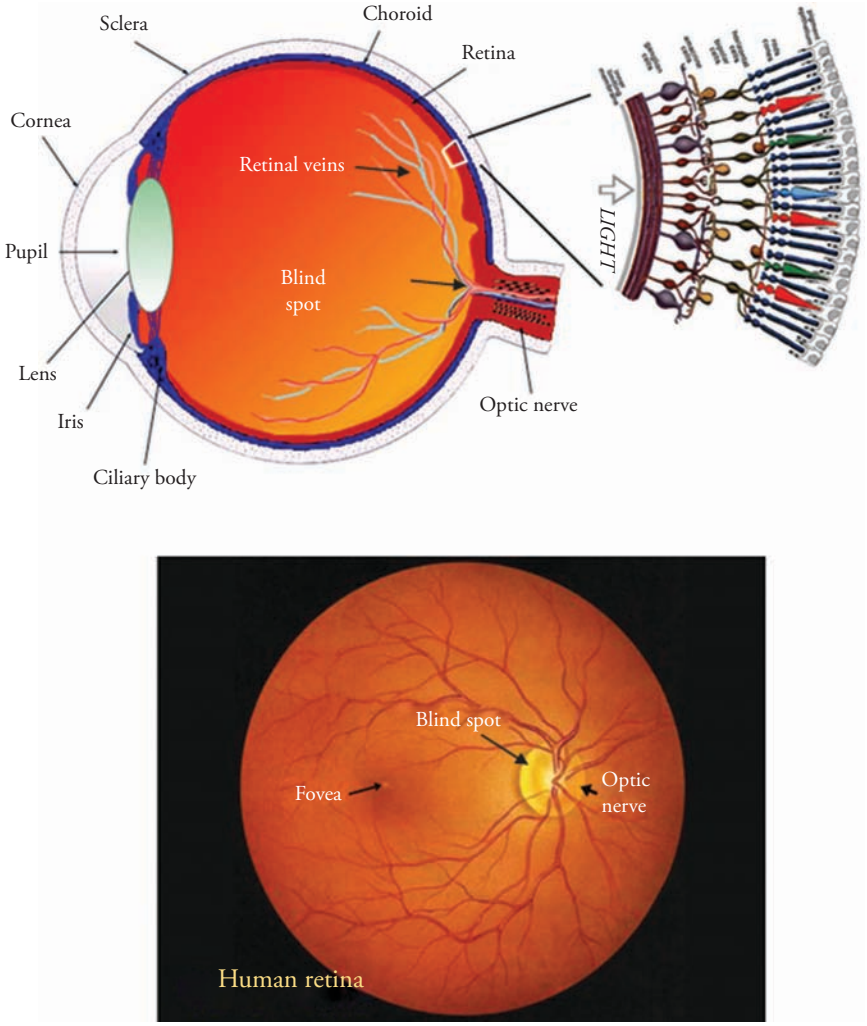


Figure I.7-5. Side and top views of the eye, showing the relationships between the fovea, blind spot, retinal veins, and optic nerve. (Images adapted with permission from Webvision-University of Utah.)

In the next few sections, we suggest how seemingly strange ideas such as "all boundaries are invisible" help to explain how the brain compensates for the noisy design of the retina. From a more general viewpoint, this discussion illustrates how deficiencies of image processing at one stage of brain organization—in this case, due to physical and developmental constraints on how the retina and optic nerve grow and how the retina is nourished by blood—are compensated by processing at higher stages of brain organization. We call this general principle *hierarchical resolution of uncertainty*. This principle helps us to understand how the noisy and rapidly metabolizing meat that makes up our eyes and brain cells give rise to the Platonic world of visual forms of which we are consciously aware as we go through our lives. It also clarifies a sense in which most of the percepts of which we are consciously aware are, mechanistically speaking, visual illusions since they reconstruct, within the brain, image properties that are occluded by the blind spot and retinal veins.

JIGGLING EYES AND OCCLUDED OBJECTS

How does the brain avoid seeing the retinal veins and the blind spot? One contributing factor is that, even when we think our eyes are fixated on a stationary object in the world, they are rapidly jiggling in their orbits, back and forth, by such a small amount that we are not consciously aware of these tiny movements. This jiggle acts like a "refresh" operation that maintains the sensitivity of visual nerve cells to the stationary object. It does so by creating small relative motions between the object and the retina. Many visual cells are sensitive to these motion transients and fire in response to them. Images that do not move relative to the jiggling retina are said to be *stabilized*. Percepts of stabilized images fade away. They cannot be seen after a little while because they are not refreshed. The retinal veins and blind spot are stabilized because they are attached to the retina. They therefore are not visible, at least not under ordinary viewing conditions.

The retinal jiggle helps to discount the veins and blind spot, but more is needed to see the world of continuous objects that we take for granted every day. To get a feel for what is missing, imagine that a simple image, like one of a thick, straight edge, is registered by the retina in a region that includes the blind spot and some retinal veins. Parts of the edge cannot effectively activate retinal photoreceptors because they are received where there are veins or the blind spot. Because the veins and blind spot are stabilized images on the retina, they fade. What remains is not a

representation of the entire edge, however, but discontinuous fragments of the edge's unoccluded regions. How does the brain transform these discontinuous fragments into a percept of a continuous edge?

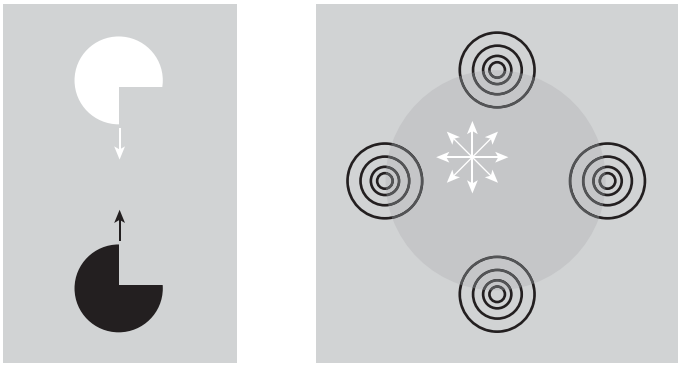
HOW ILLUSIONS MAY REPRESENT REALITY

We can now begin to understand how the brain can use the discontinuous fragments of an edge to complete a continuous edge percept. The occluded fragments of the edge are surrounded by unoccluded fragments that are approximately collinear, or aligned. The brain uses these collinear fragments to complete edge boundaries wherever they are missing. In particular, the brain has cells that can form boundaries *inwardly* between pairs, or greater numbers, of cells that are (nearly) collinear across space and whose orientational preferences are (nearly) collinear as well. In other words, a boundary can form in a region that receives no visual inputs of its own if it has collinear inducers on both sides of the region. Then, all the boundaries that were occluded by the blind spot or retinal veins can be completed. On the other hand, boundaries will not grow out of the ends of a single edge, because boundary formation takes place only inwardly between pairs or greater numbers of inducers. If these boundary-forming cells can also pool collinear signals from opposite contrast polarities, as in the case of the reverse-contrast Kanizsa square of Figure I.7-5b, then they can complete the edge even if the unoccluded edge fragments have opposite contrasts with respect to a textured background.

Using these ideas, we can begin to see how Kanizsa square percepts form in response to the images in Figure I.7-4. In particular, the brain does not known if there is a retinal vein or blind spot between the pac-men in Figures I.7-4a and I.7-4b. It will therefore complete boundaries between pairs of such collinear contrasts whenever they are long enough and not too far apart. Indeed, the existence of the Kanizsa square illusion is a small price to pay for having percepts of the world that are not occluded by the blind spot and retinal veins.

These examples illustrate why boundaries form *inwardly* between pairs, or greater numbers, of inducing contrasts. It is fortunate that boundaries cannot form *outwardly* from a single inducer. Imagine if they could: then any speck of contrast in an image or scene could force our brains to build an expanding web of boundaries that could fill our field of view. Such boundaries would do more to drive us mad than to accurately represent objects in the world. It is also clear from these examples that boundaries form in an *oriented* way between collinear, or nearly collinear, inducers, as in the case of the Kanizsa squares in Figure I.7-4a. Finally, we know from the example of the reverse-contrast Kanizsa square in Figure I.7-4b, and the Kanizsa disk-on-texture display in Figure I.7-3, that boundaries are *insensitive* to contrast polarity. These three properties of boundary completion are summarized in Figure I.7-6.

In contrast to the *inward* completion of boundaries, the filling-in process that completes percepts of visible surfaces propagates *outward* from individual sources of brightness or color. For example, the brightness and color of the unoccluded parts of an edge—after the stabilized blind spot and veins fade—flow outward into the occluded parts of the edge, thereby restoring brightness and color to the edge percept at positions that were occluded by the blind spot or retinal veins. This flow of brightness and color behaves much like a fluid that diffuses away from a source. It is called *filling-in*. Boundaries act like a dam that contains the flow of brightness and color and keeps it from flowing outside the contours of the edge. By acting like barriers, or obstructions, to the flow of brightness and color, the positions of boundaries can become visible when a different brightness or color occurs on opposite sides of the boundary. In the case of the Kanizsa square of Figure I.7-4a, for example, the four black pac-men induce local regions of enhanced brightness, called "brightness buttons," just inside the illusory square. These brightness buttons then fill-in within the square boundary to create a percept of enhanced brightness throughout the square.



Complementary Properties of Boundaries and Surfaces

Boundary Completion	Surface Filling-in
Inward	Outward
Oriented	Unoriented
Insensitive to direction-of-contrast	Sensitive to direction-of-contrast

Figure I.7-6. Complementary properties of boundary completion and surface filling-in. (Reprinted with permission from Grossberg, 2014.)

EVERY EDGE IS AN ILLUSION

To briefly summarize: a combination of inward boundary completion and outward filling-in of surface brightness and color help to complete the representation of an edge that is occluded by retinal veins or the blind spot. A remarkable implication of this discussion is that *every edge is an illusion*.

In other words, our conscious percepts of those parts of the edge that are occluded by retinal veins or the blind spot are actively constructed by the brain, just as are the more obviously illusory percepts like the Kanizsa square. The unoccluded parts of the edge are more "real" than the occluded parts, because they generate their own visual signals at the retina. But we do not know which parts are which. Both types of parts look just as real to us. Moreover, as our eyes move along the edge, these real and illusory regions may be interchanged, with no one the wiser, as different parts of the edge are occluded. In summary, even the percept of such a simple image as a real edge will often

include a series of illusory percepts that just happen to look real.

If even real percepts can be illusory, then what criteria do we use to decide in our everyday lives that a percept like the Kanizsa square is an illusion? The concepts of boundary completion and surface filling-in help to do this, especially if we keep in mind the fact that all boundaries are invisible. One possible pragmatic definition of a percept that we may tend to *call* an illusion is: *a visual illusion is a percept whose combination of boundaries and surfaces looks unfamiliar*.

For example, in the percept of the offset grating of Figure I.7-1a, there is a vertical boundary that has no visible brightness or color. We call this unfamiliar combination of boundary and surface color an illusion. On the other hand, the part of an edge that is perceived at the locations over the blind spot is believed to be a "real" percept, even though it is not due to direct visual input from the retina, because it combines boundary and surface properties in a familiar way. Indeed, many percepts that are completed over the blind spot and retinal veins seem to be real even though they are illusions in the narrow mechanistic sense of being constructed by the brain over retinal positions that transmit no contrastive inputs to the brain. The following section attempts to distinguish between the reconstructive processes that we perceive as illusions versus those that are reconstructed but interpreted to represent what is really in the world.

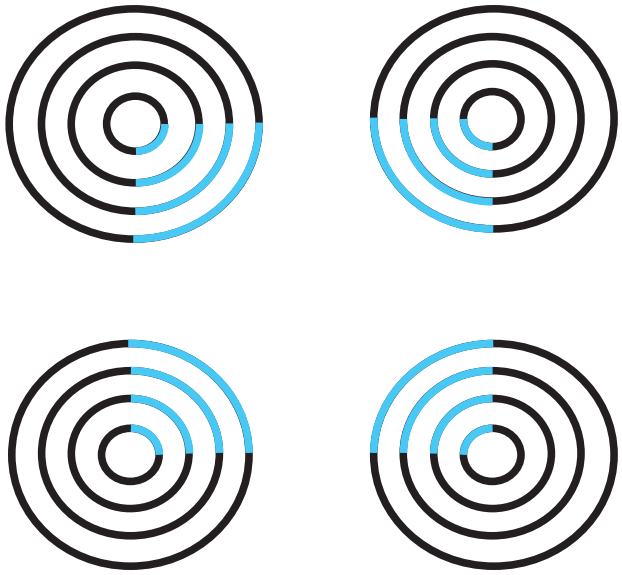


Figure I.7-7. An example of neon color spreading. Although the only blue in the figure is within the blue strips, end gaps in the boundaries that are induced by the figure where the black and blue strips intersect allow blue color to flow out of the blue strips and fill in the illusory square that is created when the end cuts at the black-to-blue edge interfaces cooperate by bipole grouping. (Reprinted with permission from Grossberg, 2008.)

EVERY LINE END IS AN ILLUSION

We noted previously that boundary completion is oriented. To initiate oriented boundary processing, the brain uses oriented detectors, called simple cells, in cortical area V1. However, such orientation-selective tuning of cells can prevent them from responding at the ends of lines. Without further processing, the open gaps in the boundaries at line ends would allow color to flow out of every line end. This additional processing can be accomplished in two stages: Spatial competition among cells at different positions that share a similar orientational preference, followed by a push-pull orientational competition among cells with different orientational preferences at the same position, with maximal competition occurring between cells sensitive to perpendicular orientations, can create line ends capable of containing the line's color within its borders. These constructed line ends are called *end cuts*.

These spatial and orientational competitive interactions are part of the process called *hierarchical resolution of uncertainty* (Grossberg, 1984, 1994; Grossberg & Mingolla, 1985b) whereby uncertainties caused by lower levels of boundary processing are compensated by processes at higher cortical levels.

NEON COLOR SPREADING

Hierarchical resolution of uncertainty can prevent color from flowing out of some line ends, but it does not prevent color from flowing out of all line ends. A famous example of color spreading from line ends is shown in Figure I.7-7. It is called *neon color spreading*. Neon color spreading illustrates key properties of color filling-in, in addition to how end cuts are formed. The image in Figure I.7-7 consists of circular annuli, part of which are black and part of which are blue. When we view this figure, we can see an illusory square filled with blue color, even though the only blue in the image is in the concentric blue circular strips.

Neon color spreading was reported in Varin (1971), who studied a "chromatic spreading" effect that was induced when viewing an image similar to the one in Figure I.7-7. Van Tuijl (1975) independently introduced images that gave rise to percepts that he called "neon-like color spreading." The following explains more fully how such illusory percepts arise from an interaction between boundary and surface system properties that are useful for survival (Grossberg & Mingolla, 1985a). For now, let us note that the black arcs and the blue arcs in Figure I.7-7 both create boundaries in our brains. At the positions where these boundaries join, the boundaries caused by the black arcs cause small breaks, called *end gaps*, to occur in the boundaries caused by the blue arcs. This happens because, in order for the percept to occur, the contrast of the black arcs with respect to the white background is chosen to be larger than the contrast of the blue arcs with respect to the white background. The boundaries formed by these contrasts are contrast-sensitive, so that the boundaries formed by the black–white contrasts are stronger than those formed by the blue–white contrasts and can thus inhibit these spatially abutting weaker boundaries using spatial competition. When these boundaries are inhibited, the cells that are tuned to other orientations, notably the perpendicular orientation, are disinhibited as part of the end-cut process. These disinhibited cells can cooperate across space to complete a boundary with the shape of an illusory square. This long-range cooperative process is often called *bipole grouping* (Grossberg, 1984). The predicted properties of bipole

grouping have been supported by psychophysical experiments (Field, Hayes, & Hess, 1993; Kellman & Shipley, 1991) and neurophysiological experiments (von der Heydt, Peterhans, & Baumgartner, 1984; Peterhans & von der Heydt, 1989) that discovered cells with these properties in cortical area V2, among other cortical areas.

Blue color can flow out of the end gaps in the broken boundaries. This filling-in of blue color across space continues until the color hits the square "illusory" boundary, which prevents its further spread. Thus boundaries act as barriers to the filling-in of brightness or color. In summary, the process of end cutting that completes line ends can also create end gaps through which brightness or color can flow when a stronger boundary abuts a collinear weaker boundary. The end cuts can, in turn, cooperate via bipole grouping to complete an illusory boundary that can contain this flow of brightness or color.

Neon color spreading illustrates how color can spread *outwardly* from the individual blue strips in all directions. Its spread is thus *unoriented* (e.g., Grossberg & Todorovic, 1988). The blue color continues to spread outwardly in all directions until it fills the illusory square boundary, beyond which it cannot spread any further. In addition to the two filling-in properties of unoriented and outward spread, we can also add the obvious property that filling-in is *sensitive* to contrast polarity, because we can consciously see its effects. These three properties of surface filling-in are summarized in Figure I.7-6.

COMPLEMENTARY PROPERTIES OF BOUNDARIES AND SURFACES

Figure I.7-6 illustrates the prediction, first made in Grossberg (1984), that the brain is organized into parallel processing streams that exhibit computationally *complementary* properties (Grossberg, 2000). Visual boundary completion and surface filling-in provide just one of many examples of brain processes with computationally complementary properties. Figure I.7-6 summarizes that boundaries are completed *inwardly* between pairs or greater numbers of inducers, in an *oriented* fashion, and are *insensitive* to contrast polarity, in the sense that they pool over opposite contrast polarities. In contrast, surface filling-in proceeds *outwardly* in an *unoriented* fashion, and is *sensitive* to contrast polarity, since surfaces can create visible percepts. These properties are manifestly complementary: inward versus outward, oriented versus unoriented, insensitive versus sensitive. They fit together like lock and key, or yin and yang.

Complementarity of boundary and surface processing is important for the success of each process. For example, filling-in needs to be unoriented so that it can cover an entire surface. On the other hand, the unoriented flow of brightness can be efficiently contained only by an oriented boundary. Likewise, a seeing process cannot efficiently build boundaries around objects in front of textured backgrounds. Both types of process are needed for either process to work well. Moreover, both types of process need to interact to overcome each other's complementary deficiencies.

The ability of boundaries to serve as barriers to filling-in provides one example of how interactions between the boundary and surface systems help them to achieve a better result than either one could attain by itself. One of the challenges facing the visual system is to combine these complementary properties into a higher synthesis, which we experience as unified visual percepts. In other words, the challenge is to reconcile the *complementarity* of boundary and surface properties with the *consistency* of conscious percepts. We call this resolution the property of *complementary consistency*.

COMPLEMENTARY PROCESSING STREAMS IN VISUAL CORTEX: BOUNDARIES AND SURFACES

How are these complementary processes represented in the brain? Much evidence suggests that they are carried out by parallel processing stream in the visual cortex. Figure I.7-8 illustrates how visual signals activate the light-sensitive retinas within our eyes. The retinas, in turn, send signals to the lateral geniculate nucleus (LGN). The LGN has a beautiful laminar organization that may, at first approximation, be broken up into a *parvocellular* part and a *magnocellular* part, so named after their parvocellular and magnocellular cell types. Output signals from the LGN branch out and activate several parallel subsystems of the visual cortex.

Three streams are easy to track in Figure I.7-8. Two of these streams proceed from the parvocellular LGN to regions of the first cortical stage, called area V1 in monkeys and area 17 in cats. One of these streams goes through

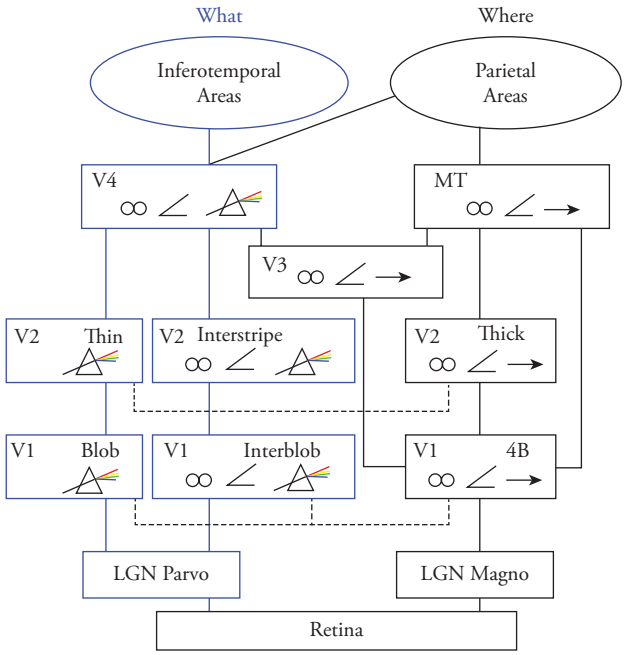


Figure I.7-8. Schematic diagram of anatomical connections and neuronal selectivities of early visual areas in the macaque monkey. LGN = lateral geniculate nucleus; V1 = striate visual cortex; V2, V3, V4, MT = prestriate cortical areas. The boundary stream goes through the blobs and thin stripes to cortical area V4 and inferotemporal areas. The surface stream goes through interblobs and interstripes to V4. The motion stream goes through V1 and MT to the parietal areas. (Reprinted with permission from DeYoe & Van Essen, 1988.)

structures that are called the *blobs* of V1. The blobs are small regions that are distributed in regular arrays throughout V1. They are highly active metabolically and therefore light up when probed by a chemical marker called cytochrome oxydase. The blobs project, in turn, to a region that is called the *thin stripes* of the prestriate cortex, in an area that is called V2 in monkeys and area 18 in cats. The thin stripes then project to prestriate area V4. Grossberg (1984) predicted that the LGN→blob→thin stripe→V4 processing stream generates visual *surface* representations and that the parallel LGN→interblob→interstripe→V4 processing stream generates visual *boundary* representations. Let us call these streams the blob and interblob streams, respectively.

Other investigators, notably Livingstone and Hubel (1984), have made related but distinct proposals. They suggested that the blob stream computes "color" and the interblob stream computes "orientation." These two proposals lead to different sets of predictions. In particular, a boundary system can complete boundaries—both "real" and "illusory"—over positions in space that receive no inputs, let alone oriented inputs; and a surface system can generate filled-in representations of figure–ground relationships that do not directly represent the local brightnesses and colors of a scene, as illustrated by the example of Kanizsa stratification in Figure I.7-2. In particular, FACADE theory predicts why completed boundaries and filled-in surfaces of the occluded parts of objects, including those in Figure I.7-2, are amodal, or invisible, even though they can have profound effects on object recognition (Grossberg, 1994, 1997). These completed representations of occluded regions receive no bottom-up oriented signals and, although they are predicted to control filling-in using double-opponent cells, generate no visible color signals. Many vision scientists now routinely use the boundary/surface distinction to interpret their experiments.

How do the complementary properties of these parallel streams arise? Grossberg (2000) predicted that they arise during brain development as part of a process of *symmetry-breaking*. This concept envisages that pairs of streams may arise from an earlier shared set of cells that bifurcate into complementary streams, while maintaining and elaborating their reciprocal connections, as developmental specialization occurs. Symmetry-breaking is a widespread phenomenon in many parts of physics and biology. In the case of the brain, current neural models suggest that, during brain development, there is a *cascade of symmetry-breaking operations on multiple spatial scales*, with some complementary processes arising as substreams of larger streams that also exhibit complementary relationships.

WATERCOLOR ILLUSION AND NEON COLOR SPREADING

Before going on to discuss other types of complementary processing during visual perception, let us consider some other examples of how complementary processing creates visual illusions of form and color. Here we focus on conceptual themes and examples rather than on mechanistic explanations of visual illusions. All of these illusions have, however, been given quantitative mechanistic neural explanations, supported by computer simulations. For readers who prefer more heuristic discussions. Grossberg & Pinna (2012) provides a heuristic summary of enough modeling concepts and mechanisms to explain a wide range of illusory percepts devised by Pinna (2009).

One of the most famous visual illusions that was studied by Pinna (1987; also see Pinna, Brelstaff, & Spillmann, 2001) is the *watercolor illusion*. The watercolor illusion is a long-range spread of color diffusing from a thin colored contour running parallel and contiguous to a darker chromatic contour. The illusion can also impart a strong figural effect across large regions. In Figure I.7-9, light blue undulating contours flanked by darker blue contours are perceived as irregular curved shapes evenly colored by a veil of light blue spreading from the inner light blue contours.

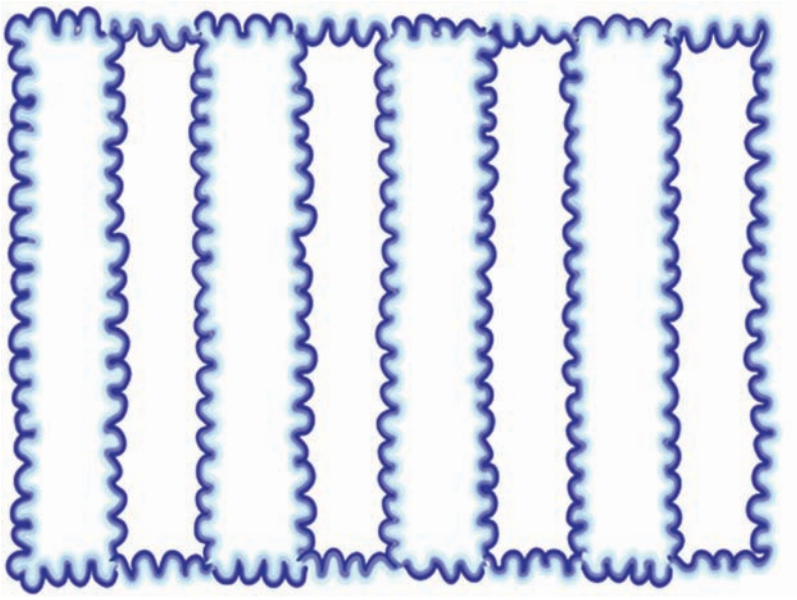


Figure I.7-9. An example of the watercolor illusion. When dark blue and light blue contours are juxtaposed, the stronger boundaries that are induced by the dark blue contours inhibit the weaker boundaries that are induced by the light blue contours, more than conversely. As a result, light blue color flows past the weaker boundaries and fills in the intervening spaces. How this creates a percept of blue figures on a white background is explained in Pinna and Grossberg (2005). (Reprinted with permission from Pinna and Grossberg, 2005.)

The regions into which the light blue spreads look like figures, whereas the remaining white regions look like background. Pinna and Grossberg (2005) explain this and other examples of the watercolor illusion. One relevant neural mechanism is the spatial competition that occurs between nearby boundaries, which also helps to explain how line ends are completed and how neon color spreads out of end gaps in boundaries (see previous discussions in this chapter). In a watercolor illusion image, just as in a neon color spreading image, one boundary is stronger than the other. The stronger boundary in this case is induced by the dark blue contours. The stronger boundary can inhibit the nearby weaker boundary of the light blue contours, more than conversely, via spatial competition. The color that the weaker boundary would otherwise contain can then partially spread, or fill in, onto the contiguous surface.

As noted earlier, the spatial competition that triggers the watercolor illusion is part of a process of hierarchical resolution of uncertainty whereby the boundary completion system can overcome uncertainties that are created by the orientationally tuned processing of edge, texture, and shading signal, thereby completing object boundaries that can usually, but not always, contain the filling-in of brightnesses and colors within the correct object regions.

TEXTURE GROUPING, BREAKING CAMOUFLAGE, AND VISUAL ART

We can now better see how visual illusions help to understand how the visual brain works. Because of their unexpected combinations of boundary and surface properties, they provide important additional information about perceptual processes that we might otherwise overlook. Second, they are crucial for our survival, as is illustrated by the fact that the percept of even of an object's edge may be, from a mechanistic perspective, "illusory." There are hundreds of examples of visual illusions that are important for our survival, including examples that depend on invisible boundaries.

Figure I.7-10 provides one such example, a famous picture of a Dalmatian dog in a snowy terrain. The white coat of the Dalmatian blends into the snowy terrain, so that the Dalmation first looks like a disconnected series of black blotches on white paper. We gradually recognize the Dalmatian when emergent boundaries form that connect the black blotches. These boundaries provide important information about the Dalmatian's shape. Although these boundaries are illusory and invisible, the brain cells that compute the boundaries can trigger signals that activate the object recognition centers in our brains. The brain uses these invisible boundaries to *recognize* the Dalmatian even if the boundaries are not themselves *seen*. The Dalmatian can hereby be recognized using its invisible boundaries in just the same way that the invisible vertical boundary of the offset grating in Figure I.7-1a can be recognized. The boundaries will only be seen if they fill-in the surface representation of the Dalmatian with a different brightness, color, or depth than its background, as in the case of the Kanizsa square (Fig. I.7-4).

Much neurophysiological evidence suggests that object recognition occurs in the *inferotemporal cortex* (e.g., Lueschow, Miller, & Desimone, 1994; Zoccolan, Kouh, Poggioa, & DiCarlo, 2007), whereas perceptual boundaries and surfaces are formed at earlier processing stages that are found in the interblob and blob processing streams of the *prestriate visual cortex*. We can recognize boundaries that we cannot see in part because recognition and seeing are carried out by different parts of the brain.

The Dalmatian example is important because it illustrates a problem that our ancestors may have confronted in the wild. Imagine that the image is not of a Dalmatian but rather of a deadly predator, like a snow leopard, that is stealthily tracking a man in the snow. All that he can see are the predator's black spots, because its white coat blends into the snow. By using his emergent invisible boundaries, he can hopefully recognize the predator in time to initiate evasive or other survival techniques. Emergent boundaries, whether invisible or not, help to



Figure I.7-10. Dalmation in snow. When this image is first seen, it may look like a textured pattern of meaningless black and white textures. After sustained inspection, amodal boundaries are completed and enable recognition of a Dalmation in snow.

break through camouflage in the many situations where local properties of an object—such as its curves and surface elements—would be insufficient for its recognition. In fact, one way to disable our ability to break through camouflage is to generate images whose orientations create emergent boundary groupings that compete with those of the target.

Figures I.7-11 through I.7-13 provide some other examples of how invisible boundary groupings can segregate textured regions from one another, or separate figures from their backgrounds. Figure I.7-11 is a famous image due to the distinguished American psychologist Jacob Beck. It is a texture composed of the letters L and T. Although the regions with vertical and diagonal T's are contiguous, the vertical T's group together with the vertical L's, rather than with the diagonal T's. One reason for this is that the vertical T's and L's share both vertical and horizontal lines, whereas the diagonal T's do not. Invisible vertical and horizontal emergent boundaries can connect, and thereby group together, the vertical T's and L's but not the diagonal T's with the vertical T's. The boundary grouping process is thus highly sensitive to the relative orientation of image contrasts. As in the Kanizsa square, inducer orientations that are *collinear*—that is, have the same orientation and are aligned across space—are strongly favored in the grouping process. Because the properties of grouping take precedence over the different identities of the T's and L's, this percept is consistent with the hypothesis that collinear grouping occurs before object recognition and cannot be overridden by top-down cognitive factors when the emergent groupings are unambiguous.

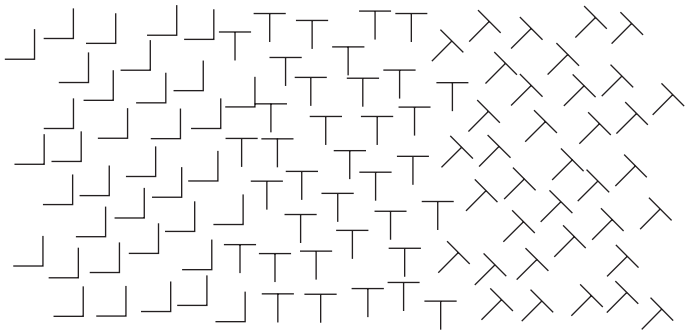


Figure I.7-11. A texture composed of T's and L's. The vertical T's in the middle group with the vertical L's on the left, rather than the oblique T's on the right, thereby illustrating that, when boundary grouping can occur unambiguously, it cannot be overridden by higher-level recognition processes. (Reprinted with permission from Beck, Prazdny, & Rosenfeld, 1983.)

Figure I.7-12 illustrates that emergent boundary groupings are sensitive to both the distance and the relative

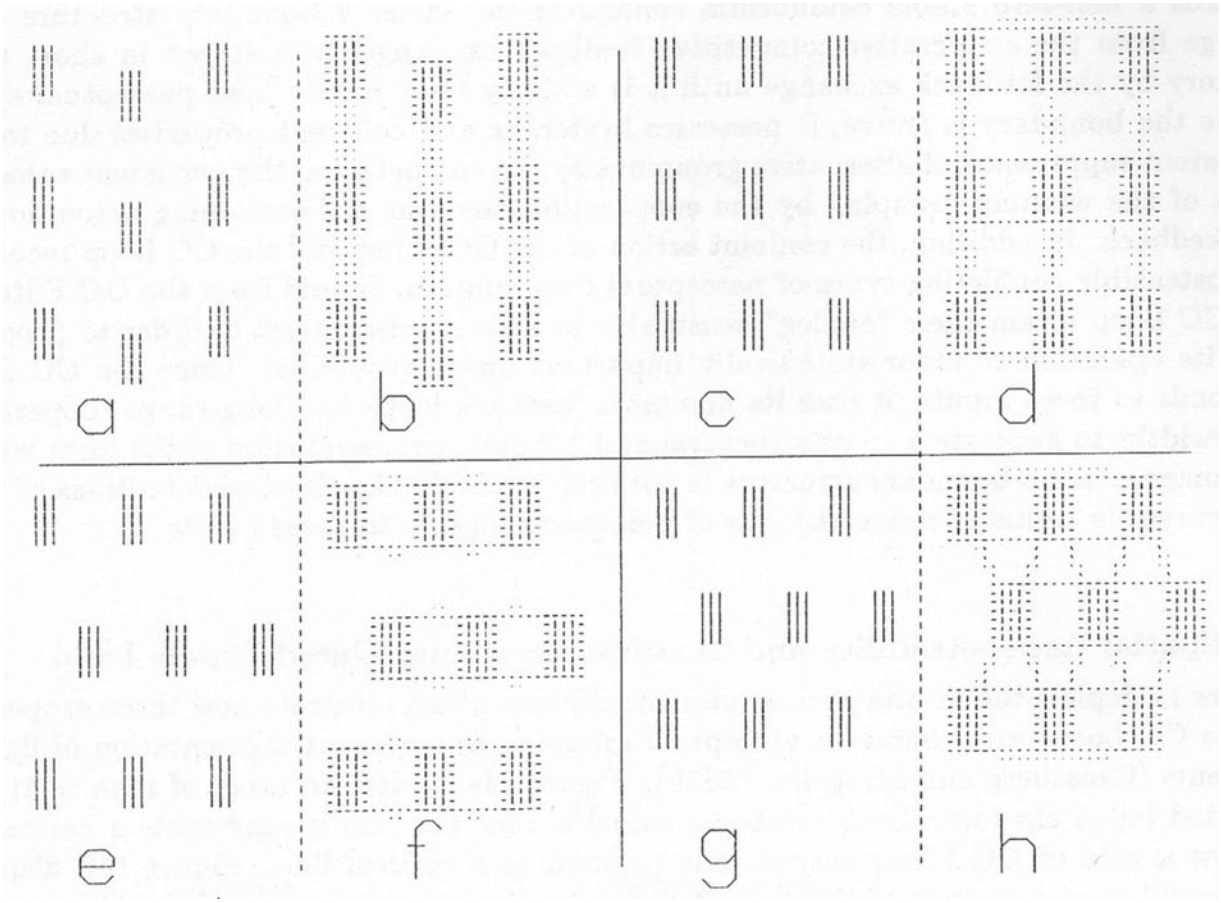


Figure I.7-12. Examples of how emergent perceptual groupings can form with orientations that are not in the inducing pictorial elements. Figures I-7-12a, I.7-12c, I.7-12e, and I.7-12g are inputs. Figures I.7-I.7-12b, I.7-12d, I.7-12f, and I.7-12h are model simulations of the emergent groupings. (Reprinted with permission from Grossberg & Mingolla, 1987.)

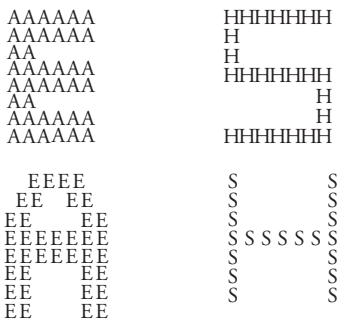


Figure I.7-13. Emergent boundary groupings can generate recognizable letters from local image inducers that do not have the orientations of the letters.

orientations between the inducing elements. It is easier for boundaries to form between elements that are closer together, other things being equal. The importance of such "Gestalt rules" for visual perception was advocated by German psychologists such as Wolfgang Köhler and Max Wertheimer during the first half of the 20th century. Although each of the four images in Figure I.7-12 is composed of three rows of six vertical line segments, their spatial arrangement can generate strikingly different percepts. The percept in Figure I.7-12a is of emergent vertical boundaries, the one in Figure I.7-12c is one of both emergent vertical and horizontal boundaries, the one in Figure I.7-12e is of emergent horizontal boundaries, and the one in Figure I.7-12g can generate both emergent horizontal and diagonal boundaries. Percepts of horizontal and diagonal boundaries emerge in response to Figures I.7-12c, I.7-12e, and I.7-12g images even though there are only vertical line segments in the inducing images. The global arrangement of the line elements determines which grouping will win. Figures I.7-12b, I.7-12d, I.7-12f, and I.7-12g show model simulations of these emergent boundary groupings. These boundary properties use the same combination of spatial competition, orientational competition, and bipole grouping that were used to explain neon color spreading.

One implication of this result can immediately be illustrated in Figure I.7-13, where the boundaries of the letters E, S, A, and H are generated as emergent properties of groupings of the letters A, H, E, and S, respectively. This sort of global "forest before trees" structuring of a percept was particularly emphasized by David Navon. In the vertical parts of the global S, collinear grouping of local orientations is used, whereas in the horizontal parts of the global S, groupings perpendicular to the local orientations form the emergent boundary. The top of the global letter A, in contrast, uses oblique groupings, even though only local horizontal and vertical orientations occur in its constituent letters E.

Many percepts that we observe in real life are illusions of this type. We are continually separating figures from their backgrounds based on boundary groupings that form between regions with different textures and colors. Many visual artists, notably the Impressionists and Fauvists, exploited this property of visual perception to create some of the most beautiful paintings ever made. They were masters at showing how to arrange small patches of color on a canvas in such a way that the groupings that were thereby induced in the brain can be recognized as objects in the world.

In particular, many artists instinctively exploited the fact that invisible boundaries indirectly assure their own visibility through their interactions with the surface stream. Within the surface stream, boundaries trigger the filling-in of lightness and color within their contours and restrict the spreading of lightness and color to the surface region that they bound. In addition, the same visual stimuli activate both the boundary and the surface streams. Thus a visual scene or painting activates both the boundary and the surface streams in parallel, the boundary stream generates an emergent perceptual grouping, and this grouping controls the filling-in of potentially visible surface lightnesses and colors within the surface stream. These facts about brain processing have influenced the creative struggles and decisions of many artists.

Every artist faces the choice of whether to draw edges around the surfaces that his or her painting will represent, or to allow the brain to create boundaries through the process of amodal perceptual grouping. Drawing in an edge with a particular color may influence the appearance of nearby colors either through color assimilation, which may darken the nearby colors, or by color contrast. Matisse (1947/1992), among others, struggled with this issue for many years. Later in his life, when his health led him to work with paper cut-outs, he wrote about "the external conflict between drawing and color . . . Instead of drawing an outline and filling in the color . . . I am drawing directly in color." Matisse was already "drawing directly in color" in his paintings from the Fauve period, as illustrated in his painting from 1905 called *The Roofs of Collioure* (Fig. I.7-14). Matisse realized instinctively that, if he painted directly with appropriately shaped color patches, these patches would induce the formation of amodal boundaries within the brain of the viewer. These boundaries, in turn, would capture the inducing colors to form the surface representations of color and form that enable the viewer to understand the painting, but without the risk that edges with colors of their own could darken the perceived surface colors (Fig. I.7-15). These concepts clarify the sense in which a masterpiece like *The Roofs of Collioure* is a visual illusion. Grossberg (2008) analyzes various artistic theories and struggles of famous artists from the perspective of neural models of boundary and surface formation.

FROM 2D TO 3D: EFFECTS OF CONTEXT ON 3D BOUNDARY ORIENTATION AND SURFACE FILLING-IN

The previous summary includes discussions of how figure–ground illusions can arise in which partially occluded objects can be completed behind their occluders (e.g., Figs. I.7-2 and I.7-4). However, this barely grazes the huge field of 3D vision and figure–ground perception, which goes beyond the scope of this chapter, although a great deal of neural modeling work in the FACADE and 3D LAMINART theories has clarified how it works and explained many perceptual and neurobiological experiments along the way (e.g., Cao & Grossberg, 2005, 2012; Fang & Grossberg, 2009; Grossberg, 1987, 1995, 1997; Grossberg & Howe, 2003; Grossberg & McLoughlin, 1997; Grossberg



Figure I.7-14. *The Roofs of Collioure* was painted by Matisse in 1905.

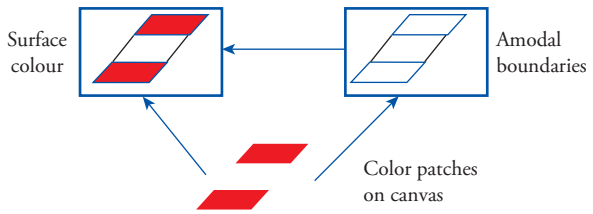


Figure I.7-15. *The Roofs of Collioure* in Figure I.7-14 illustrates how Matisse "paints in color" without using explicitly drawn edges that could darken the painting's perceived colors. These color patches generate amodal boundaries that can capture the surface colors into surface percepts that are meaningful to viewers. (Reprinted with permission from Grossberg, 2008.)

& Swaminathan, 2004; Grossberg & Yazdanbakhsh, 2005). Here, only one main point is noted using Figures I.7-16 and I.7-17 to illustrate how 3D illusions can occur.

The top image in Figure I.7-16 shows how a rectangular parallelopiped can be constructed in a 2D picture from sides that individually generate flat percepts of rectangles and parallelograms. Consider, in particular, the flat-looking parallelogram in the middle of the figure. If it is surrounded by sides as in the left-hand parallelopiped, then its lower side seems to project from near to far. However, if it is surrounded by sides as in the right-hand parallelopiped, then the same parallelogram's lower side appears to project from far to near. This simple demonstration shows that the percepts of the parallelopipeds are highly context-sensitive and depend on the way in which the sides are joined together at their shared edges.

When both types of parallelopipeds are joined together in the famous Necker cube at the bottom of the figure, then the 3D interpretations of both parallelopipeds can alternate through time. Focusing attention on one of the edges can, moreover, pull it forward and cause the side to which it belongs to look closer in space.

Figure I.7-17, which is due to Peter Tse, shows in a striking way that lines that are perceived as parallel when interpreted in 2D no longer look parallel when interpreted as part of a 3D percept, even though the entire figure is a 2D picture.

Grossberg and Swaminathan (2004) explain and quantitatively simulate such percepts, including the bistable 3D surface percepts of the Necker cube as they oscillate through time. In order to explain such percepts, the concept of bipole grouping was generalized from 2D to 3D, and the fact that it could be so generalized is another testimonial to its existence in the brain. In particular, Grossberg and Swaminathan (2004) show through computer simulations how bipole grouping cells can develop in response to visual inputs to represent collinear 2D and 3D groupings. The latter cells, called *disparity gradient cells*, can represent tilted edges such as those that occur in the percepts of the parallelopipeds. The same developmental laws gives rise to *angle cells* that are sensitive to the angular relationships that occur when two or more edges terminate at the same position in space. Interactions between angle cells, disparity gradient cells,

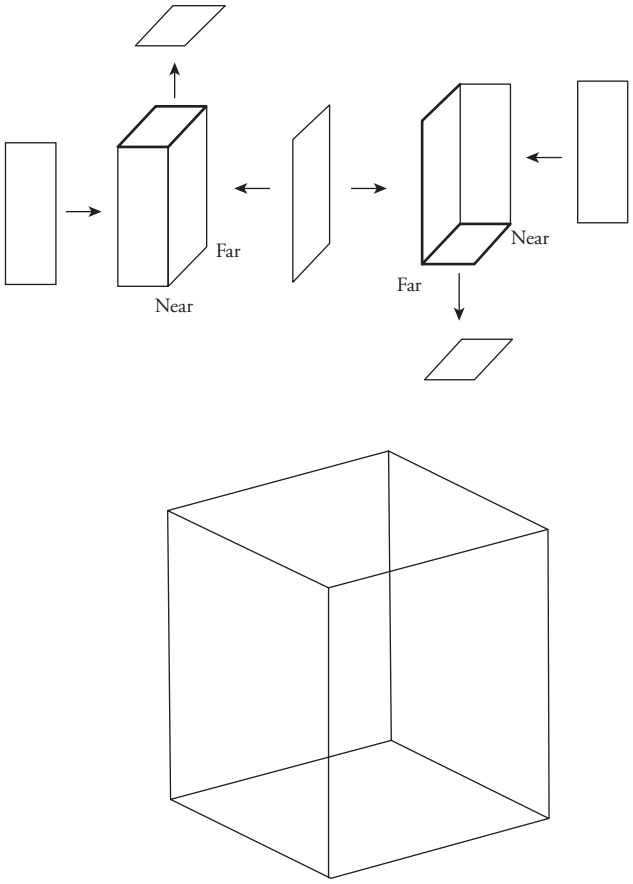


Figure I.7-16. The top figure shows how flat rectangles and parallelograms can generate 3D slanted percepts of parallelopipeds. (Reprinted with permission from Grossberg & Swaminathan, 2004.) The bottom figure of the Necker cube generates a bistable percept through time of both 3D interpretations of these parallelopipeds. (Reprinted with permission from Necker, L.A. (1832). "Observations on some remarkable optical phaenomena seen in Switzerland; and on an optical phaenomenon which occurs on viewing a figure of a crystal or geometrical solid". *London and Edinburgh Philosophical Magazine and Journal of Science.* **1** (5): 329–337.)

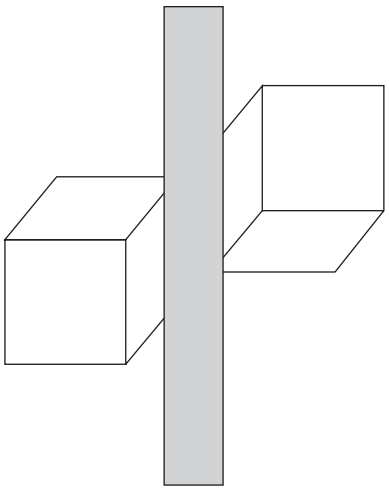


Figure I.7-17. Although the diagonal lines that are partially occluded by the vertical gray bar are colinear in 2D, they are not perceived to colinear as part of their respective parallelopipeds. (Reprinted with permission from Tse, 1999.)

and bipole cells give rise to the context-sensitive boundaries of the parallelopiped percepts. In addition, Grossberg and Swaminathan explain and simulate how these 3D boundaries capture surface filling-in that can occur over surfaces that are tilted in depth.

COMPLEMENTARY PROCESSING STREAMS IN VISUAL CORTEX: FORM, MOTION, AND FORMOTION

The organization of the visual cortex into parallel processing streams with complementary properties, and the way in which interactions across the streams generates consistent percepts, helps to explain many visual illusions. These complementary relationships may occur both within and across the "What" and "Where" visual cortical streams (Grossberg, 2000). For example, neural models explain perceptual and brain data about interactions of visual form and motion by explicating their complementary mechanisms (Baloch & Grossberg, 1997; Berzhanskaya, Grossberg, & Mingolla, 2007; Francis & Grossberg, 1996b), Object tracking and spatial navigation also use complementary mechanisms (Browning, Grossberg, & Mingolla, 2009a, 2009b; Grossberg, Mingolla, & Pack, 1999; Pack, Grossberg, & Mingolla, 2001).

The pale stripes of cortical area V2, also called interstripes, carry out 3D boundary grouping of object form, whereas the middle temporal, or MT, cortical area carries out motion processing. Grossberg (1991) predicted that the form and motion processing streams exhibit complementary properties due to the conflicting computational requirements of computing depth-selective object form versus those of computing directionally-selective object motion. According to this view, by itself, V2 can compute good estimates of object depth but only coarse estimates of object motion direction, whereas MT can compute good estimates of object motion direction but only coarse estimates of object depth. Form-to-motion interactions from V2 to MT, also called *formotion* interactions, were predicted to select motion direction signals in MT with the correct depths, and thus to enable good object tracking in depth to be carried out by subsequent processing stages of the Where cortical processing stream. This Grossberg (1991) prediction has been embodied in a model called the 3D FORMOTION model. The formotion prediction differs from that of Livingstone and Hubel (1988) and DeAngelis, Cumming, and Newsome (1998), who propose that precise depth estimates are created in MT. The Grossberg prediction has been supported by neurophysiological experiments of Ponce, Lomber, and Born (2008), who cooled V2, thereby eliminating the V2-to-MT influence, and recorded from MT to find the predicted good estimates of object motion but coarse estimates of object depth.

Formotion interactions have been used to explain many visual illusions. For example, in Grossberg (1991), they were used to explain the apparent motion of subjective surfaces (Fig. I.7-18) that was reported by Ramachandran, Rao, and Vidyasagar (1973) and Ramachandran (1985). To generate this percept, two visual frames alternate through time. In one frame, there are four pac-man figures that induce a Kanizsa square percept. To the right of the pac-men is a

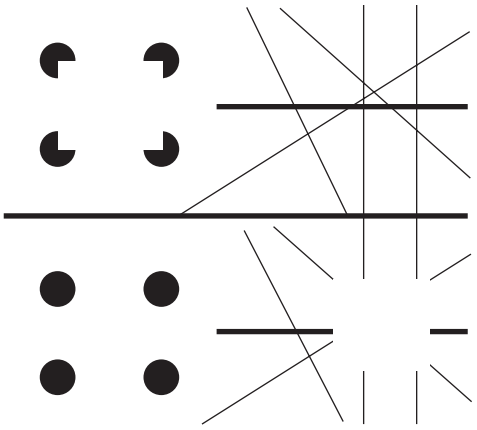


Figure I.7-18. Apparent motion of illusory contours. These images demonstrate that apparent motion of illusory contours arises through the interaction of the static illusory contours via formotion interactions that are predicted in Grossberg (1991) to occur from cortical area V2 to cortical area MT. Frame 1 (row 1) is temporally followed by Frame 2 (row 2) and conversely. (Reprinted with permission from Ramachandran, 1985.)

random array of intersecting oriented lines. In the second frame, the pac-men are replaced by four filled circles and the lines are deleted in the region of an imagined square, again producing an illusory square percept. By this construction, there are no features that can be matched across the two figures. Despite this fact, when the frames alternate through time, a percept of a square moving from left to right and back again is clearly perceived. Mechanisms of boundary completion and surface filling-in were used to explain the illusory square percepts. Mechanisms of long-range apparent motion (Grossberg & Rudd, 1992) were used to explain the motion percept in response to boundary signals in MT that were received from V2. Many other visual illusions have also been explained using these formotion mechanisms, including the line motion illusion, motion induction, and transformational apparent motion (Baloch & Grossberg, 1997).

COMPLEMENTARY PROCESSING STREAMS IN VISUAL CORTEX: MOTION PERCEPTION AND TRACKING

Area MT is broken into substreams (Born & Tootell, 1992): MT^- is involved in generating percepts of object motion and the tracking of objects moving relative to an observer, whereas MT^+ supports visually-based navigation as an observer moves with respect to the world. MT^- inputs to the ventral medial superior temporal (MSTv) cortex, whereas MT^+ inputs to the dorsal medial superior temporal (MSTd) cortex. These substreams obey computationally complementary subtractive versus additive laws, respectively, and interact to overcome their complementary deficiencies. How these complementary substreams interact to carry out object tracking and visual navigation is modeled by Browning et al. (2009a, 2009b), Elder, Grossberg, and Mingolla (2009), and Pack et al. (2001).

One example of the functional role of these interstream interactions occurs during the maintenance of smooth pursuit eye movements (SPEM). Such movements use a negative feedback system whereby the oculomotor system continuously attempts to match the velocity of the eye to that of the object being tracked. This process takes place in the MT^-–MSTv stream. What happens when this system stabilizes the target near the fovea? At these times, there is little or no motion of the target on the retina. Therefore, the pursuit system cannot rely wholly on retinal target velocity to maintain an accurate SPEM. A number of additional signals have been hypothesized to guide pursuit, including target position, target acceleration (Lisberger, Morris, & Tychsen, 1987) and a "memory" of target velocity (Young, Forster, & van Houtte, 1968), which is often described in terms of an oculomotor efference copy of eye velocity (von Holst, 1954). An efference copy duplicates the neural signal sent to the muscles that move the eye and carries information about the movement of the eye that is independent of the retinal image. An efference copy can hereby maintain a prediction of pursuit velocity from moment to moment. Thus the brain may combine retinal information about target motion or position with extraretinal information about the velocity of eye rotation.

An additional source of information relevant to pursuit maintenance is the existence of the visual background. A SPEM is typically made across a visual scene that contains stationary objects, and these objects sweep across the retinal image with a velocity opposite that of the target. This results in large-field coherent motion across the retina. Such large-field processing goes on in the MT^+–MSTd stream. It can trigger an involuntary eye rotation known as optokinetic nystagmus (OKN). OKN causes the eye to move in the same direction as the large stimulus, so that an OKN movement to track retinal motion of a visual background during pursuit would be in the opposite direction of the ongoing pursuit movement. As such, it is crucial that optokinetic signals be suppressed during execution of a SPEM.

The visual motion of the background can provide information about the velocity of ongoing SPEMs, even when the pursuit target is relatively stable on the retina. Such a signal could also be used to generate a prediction of target velocity. The visual motion of the background therefore has contradictory effects on the pursuit system, providing a potentially useful signal for pursuit maintenance and a potentially destructive OKN signal. Pack et al. (2001) have modeled how feedback from the MT^+–MSTd stream to the MT^-–MSTv stream can be used to maintain accurate tracking speed during periods when the moving target is successfully fixated.

The complementary organization of motion tracking and spatial navigation also leads to visual illusions. For example, during an accurate SPEM, target motion on the retina is very small, while objects in the background move across the retina. Psychophysical experiments indicate that human subjects are able to estimate the velocity of objects during a SPEM, but that this ability is somewhat limited. Specifically, observers underestimate the velocity of a moving target during a SPEM when no visible

background is present (the Aubert-Fleischl phenomenon; Aubert, 1886) and perceive slight motion of a stationary visual background during a SPEM (the Filehne illusion; Filehne, 1922). Pack et al. (2001) simulate these illusions as consequences of the above interactions.

ALL OBJECT MOTION IS ILLUSORY: THE APERTURE PROBLEM

Essentially all percepts of moving objects are illusions, in the mechanistic sense that they all involve active reorganization and reconstruction of the visual signals received by the retina. For example, imagine a predator or prey darting intermittently behind protective cover in a forest. As the animal moves, patterns of light and shade play upon its body through the overhanging foliage. These moving regions mingle with the movements caused on the animal's body as its limbs and muscles deform the textured patterns on its coat. The luminance and color contours of these moving regions may move in a variety of directions that do not necessarily point in the direction of the animal's physical movement. Rather, a scintillating mosaic of moving contours may be generated that could easily prevent its detection, as in camouflage.

Detecting an animal under these rather typical conditions in the wild is obviously of great ecological importance to an observer, whether predator or prey. Doing so, however, requires that the visual system of the observer solve several basic perceptual problems, such as: How does the observer separate the forest cover from the moving animal, so that cover and animal can be independently recognized? How does the cover pop out in front of the animal so that the contours of the animal are not confused with those of the cover? How is the scintillating mosaic of moving contours reorganized into a coherent object percept with a unitary direction and speed of object motion? How are the intermittent appearances of the animal integrated into a continuous motion percept that is amodally completed behind the cover, thereby enabling the animal to more easily be tracked? How does this amodally completed motion percept adapt itself to the animal's variable speed of locomotion? How does attention track the animal's moving trajectory?

Grossberg (1998) reviews and unifies neural models that simulate how all of these processes work to generate motion percepts (e.g., Berzhanskaya et al., 2007; Chey, Grossberg, & Mingolla, 1997; Grossberg, 1994; Grossberg, Mingolla, & Viswanathan, 2001; Grossberg & Rudd, 1992; Kelly & Grossberg, 2000). This chapter briefly discusses two of these types of percepts because they are both visual illusions, at least from a mechanistic perspective, albeit illusions with high survival value: the percept of a continuous trajectory of the animal's movement behind the occluding cover, and the percept of the direction and speed of the animal's motion.

The process whereby intermittent views, or flashes, of the animal moving behind the forest cover are continuously interpolated as it moves at variable speeds is a type of long-range apparent motion. This process may use form-motion interactions to group the individual target "flashes" that are intermittently seen by the observer before these groupings are interpolated by the motion system. This sort of percept is illustrated in Figure I.7-18. Rather than being a laboratory curiosity, such long-range apparent motion is used to help track a target that is intermittently occluded.

The percept of object direction and speed can use a relatively small number of unambiguous and correct motion signals, called *feature tracking* signals, to "capture" compatible motion signals across the entire object and to suppress a much larger number of incompatible motion signals. In particular, the form boundaries of the moving animal's outer contours may generate feature tracking signals, which propagate to the ambiguous scintillating mosaic within these contours to capture compatible motion signals and suppress incompatible ones. This is a visual illusion that transforms the actual motion inputs received from the animal into a different motion percept. Doing so requires a solution of one of the most important problems facing the motion system, namely, the *aperture problem*.

Wallach (1935; see also Wuerger, Shapley, & Rubin, 1996) further analyzed the observation of Stumpf in 1911 that the motion of a featureless line seen within a circular aperture is perceptually ambiguous: Given any real direction of motion of the line, its perceived direction is perpendicular to its orientation. This phenomenon was called the aperture problem by Marr and Ullman (1981). The aperture problem is faced by any localized neural motion sensor, such as a neuron in the early visual pathway, which responds to a moving local contour through an aperture-like receptive field. Only when the contour within an aperture contains features, such as line terminators, object corners, or high contrast blobs or dots, can a local motion detector accurately measure the direction and velocity of motion (Shimojo, Silverman, & Nakayama, 1989). For example, when the aperture is rectangular, as during the barberpole illusion (Wallach, 1935), moving lines may appear to move in the direction parallel to the longer edges of the rectangle within which they move, even if their actual motion direction is not parallel to these edges. The brain must solve the aperture problem, despite the fact that every neuron's receptive field defines an "aperture," in order to detect the correct directions of moving objects in the world. The examples of circular and rectangular apertures, or occluders, provide important cues about how the brain often succeeds in doing this in the real world.

When an object moves behind multiple occluders, aperture ambiguities can again lead to nonveridal percepts of its real motion. Despite the fact that the object may be segmented into many visible parts by the occluders, the visual system can often integrate these parts into a percept of coherent object motion that crosses the occluders. Studying conditions under which the visual system can and cannot accomplish this provides important cues to the processes that are used by the visual system to create useful and predictive object motion percepts during normal viewing conditions.

The aperture problem shows that, under many viewing conditions, the visual system perceives an illusory direction of motion that is perpendicular to a contour's orientation. The brain tries to solve the aperture problem by transforming one illusion into another. The latter illusion, however, helps to predict the direction and speed in which objects really move.

As noted earlier, the visual system uses the relatively few unambiguous motion signals arising from image features, called feature tracking signals, to select the ambiguous motion signals that are consistent with them, while suppressing the more numerous ambiguous signals that are inconsistent with them. For example, during the barberpole illusion, feature tracking signals from the moving line ends along the longer edges of the bounding rectangle of the barberpole compute an unambiguous motion direction. These feature tracking signals can become strong in the brain's motion processing stream within cortical areas V1, MT, and MST because directionally-selective cells can accumulate consistent evidence for that motion direction through time. In contrast, within the interior of a moving line, multiple directions become partially activated, and these compete using a self-normalizing competition, thereby weakening each other.

Later feedback processes between MT and MST enable these sparse, but relatively strong, feature tracking signals to gradually propagate through time into the interior of the rectangle. This *motion capture* process selects the feature tracking motion direction from the ambiguous directions along the lines within the rectangle. Motion capture also suppresses the other ambiguous motion signals (Ben-Av & Shiffrar, 1995; Bowns, 1996, 2001; Castet, Lorenceau, Shiffrar, & Bonnet, 1993; Chey et al., 1997; Chey, Grossberg, & Mingolla, 1998; Grossberg et al., 2001; Lorenceau & Gorea, 1989). When a scene does not contain any unambiguous motion signals, the ambiguous motion signals cooperate to compute a consistent object motion direction and speed (Grossberg et al., 2001; Lorenceau & Shiffrar, 1992). Chey et al. (1997) simulated the time course of such motion capture, and Pack and Born (2001) described neurophysiological data from cortical area MT of monkeys that fit this predicted time course.

MOTION CAPTURE AND THE TEMPORAL DYNAMICS OF DECISION-MAKING

The way in which the 3D FORMOTION model solves the aperture problem through motion capture helps to explain many visual illusions of motion. In particular, it sheds new light on how the brain makes movement decisions by selecting the direction of saccadic eye movements to match the direction of probabilistically defined motion stimuli (Grossberg & Pilly, 2008). This type of movement decision is based on a visual illusion whereby the brain converts an image that consists mostly of randomly moving dots, with a subset of dots briefly moving in a fixed direction, into an estimate of a coherent direction of dot motion. This conversion seems, once again, to be due to the brain's mechanisms for solving the aperture problem.

It is well known that speed and accuracy of perceptual decisions covary with certainty in the input, and correlate with the rate of evidence accumulation in parietal and frontal cortical neurons. The 3D FORMOTION model creates an emerging estimate of the direction of motion as motion capture sets in. This estimate needs to be converted into an estimate of directional eye movement. The Motion Decision (MODE) model (Fig. I.7-19) adds to 3D FORMOTION processing a subsequent processing stage for making the coordinate change from visual motion direction to eye movement direction. This subsequent stage is interpreted to occur in the lateral interparietal region (LIP) of the parietal cortex, where a number of neurophysiologists have recorded cells whose dynamics covary with the behaviorally observed properties of saccadic eye movements in response to the moving dots (e.g., Roitman & Shadlen, 2002; Shadlen & Newsome, 2001).

This LIP circuit is modeled using a kind of decision circuit that has become classical in the neural modeling literature, namely, a network of cells that obey the membrane equation, or shunting, dynamics that are known to occur in neurons, when they interact via a recurrent on-center off-surround network. Such a network is often called a recurrent competitive field (RCF), and its variants have been used by a number of authors to model the dynamics of perceptual or motor decisions in both deterministic models (e.g., Brown, Bullock, & Grossberg, 2004; Chey et al., 1997; Francis, Grossberg, & Mingolla, 1994; Francis & Grossberg, 1996a, 1996b) and stochastic models (Boardman, Grossberg, Myers, & Cohen, 1999; Cisek, 2006; Grossberg, Roberts, Aguilar, & Bullock, 1997; Usher & McClelland, 2001). Grossberg (1973) was the first to mathematically define and analyze these networks, proving along the way how they can make winner-take-all and distributed decisions. See Grossberg (1980) for a heuristic review of these early results. In order to explain the motion decision-making data, the RCF network needs to have its outputs gated by the basal ganglia. How the basal ganglia can gate the release of saccadic eye movement outputs was modeled in detail by Brown, Bullock, and Grossberg (1999, 2004). Cisek, Puskas, and El-Murr (2009) have also modeled how the buildup of evidence for movement decisions depends on basal ganglia gating.

Newsome, Shadlen, and colleagues studied neural correlates of perceptual decision-making in macaque monkeys who were trained to discriminate motion direction. Random dot motion displays covered a 5^{o} diameter aperture centered at the fixation point. Motion coherence was controlled by varying the fraction of dots moving nonrandomly in a particular direction from one frame to the next in each of three interleaved sequences. Varying motion coherence provided a quantitative way to control the ambiguity of directional information that the monkey used to make a saccadic eye movement to a peripheral choice target in the perceived motion direction, and thus the task difficulty. Apart from being few in number, these correct directional signals also have a short life span because a new set of signal dots are chosen in every frame.

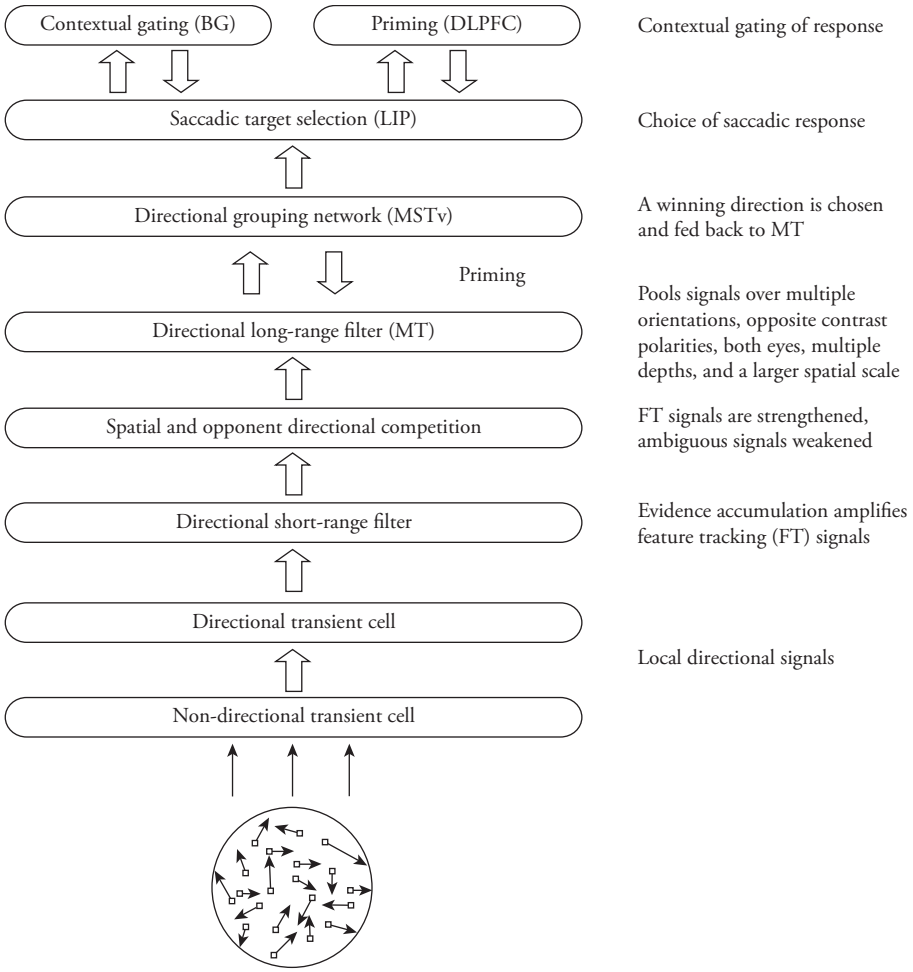


Figure I.7-19. MODE model processing stages retina/LGN–V1–MT–MST–LIP–BG. (Reprinted with permission from Grossberg & Pilly, 2008.)

Neurophysiological recordings were done in LIP while the monkeys performed these tasks. Figures I.7-20 though I.7-22 summarize some of the behavioral and LIP neurophysiological data, along with the MODE model simulations of them. These simulations are included to illustrate the power of current neural models for quantitatively explaining and predicting perceptual and brain data, including data about visual illusions. Note that, despite the discrete and probabilistic nature of the input stimuli, both the behavioral and neurophysiological data change continuously in time, reflecting the growing coherence of the neuronal responses due to motion capture interactions.

CONTEXT CAN CREATE AMBIGUITY AND THE INFORMATION NEEDED TO OVERCOME IT

Why does a solution of the aperture problem help to solve a movement decision problem in response to moving dots, given that individual dots have unambiguous motion directions and thus can generate feature tracking signals? Why is motion capture needed when many dots move in the same image? A moment's thought clarifies that a similar question needs to be asked when there are multiple moving elements in any scene. Our answer to this question illustrates how context can create computational ambiguities, as well as provide information that brain interactions can use to overcome them, and thereby generate "illusory" perceptual representations that are more informative than local cues could ever be.

The directional filters that accumulate directional evidence to generate feature tracking signals in response to a single dot can be fooled when there are multiple dots in each frame and some dots move incoherently. These directional filters can generate local directional signals between *any* two dots that occur with an appropriate spatiotemporal displacement. Typically some directions will be amplified more than others, but lower motion coherence, higher dot density, and more interleaving of stimulus frames increase the probability that incorrect directional signals will be generated, thereby reducing the impact of correct local groupings in determining a clear motion directional percept. In addition, apart from being few in number, these correct directional signals also have a short life span because a new set of signal dots are chosen every frame.

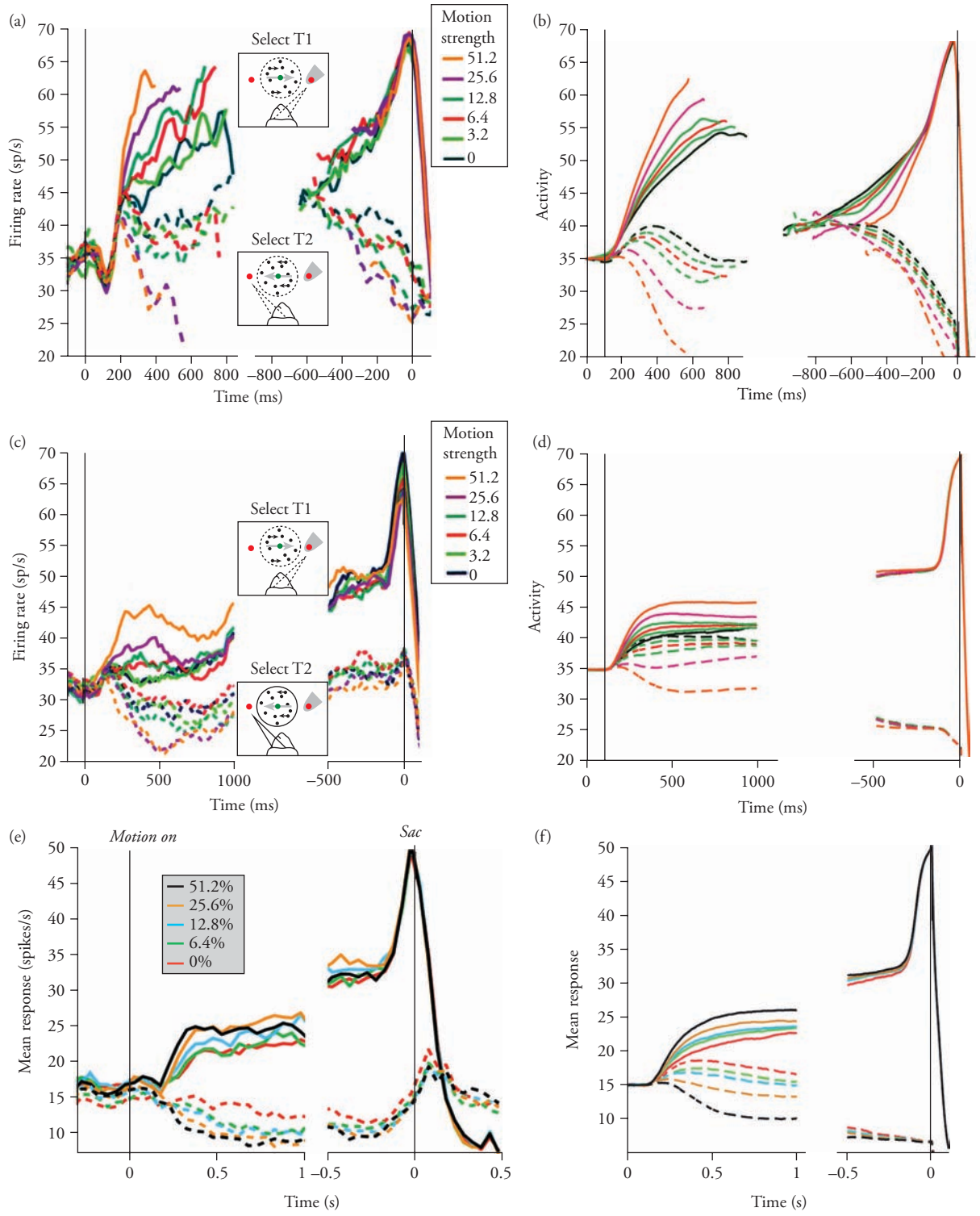


Figure I.7-20. Temporal dynamics of LIP neuronal responses during fixed duration (FD) and reaction time (RT) tasks. (a) Average responses of a population of LIP neurons among correct trials during the RT task. (b) Model simulations replicate LIP cell recordings during the RT task. (c) Average responses of a population of LIP neurons among correct trials during the 2002 FD task. (d) Model simulations mimic LIP cell recordings during the FD task. (e) Average responses of a population of LIP neurons among correct trials during another FD task. (Reprinted with permission from Grossberg & Pilly, 2008.)

Therefore, the motion stream must somehow enable a relatively sparse set of short-lived and correct feature tracking signals to gradually discount the more numerous incorrect local directional groupings. The motion capture process that is predicted to occur as a result of motion grouping feedback signals between MT and MST can do this, with the results shown in Figures I.7-18 through I.7-20.

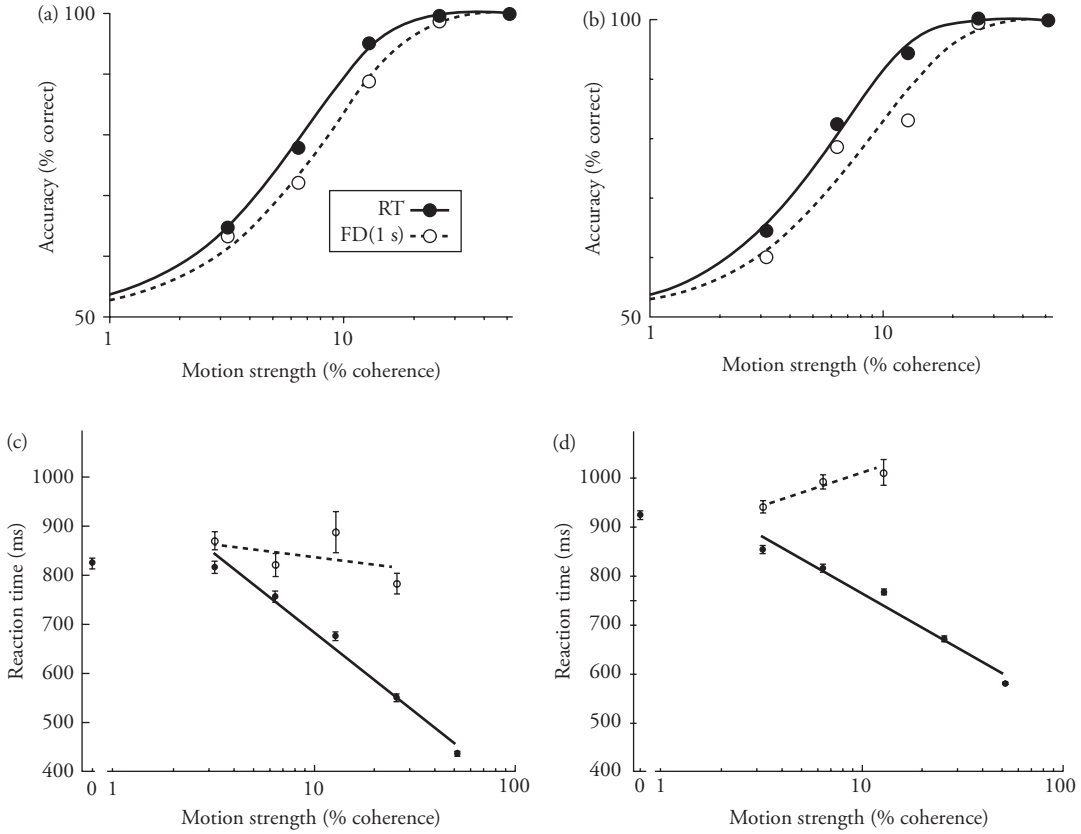


Figure I.7-21. Psychometric and chronometric data during the FD and RT tasks. (a) Accuracy data (percentage correct) as a function of motion coherence. (b) Model simulations emulate these data. (c) Reaction time data (ms) as a function of motion coherence (percent certainty). (d) Model simulations emulate the data on both correct and error trials. In (a–d), the abscissa is in the $\log_{10}$ scale. (Reprinted with permission from Grossberg & Pilly, 2008.)

HOW DOES THE BRAIN COPE WITH A PROBABILISTIC WORLD?

Earlier we mentioned the contributions of Kanizsa to better understanding how primarily bottom-up visual processes contribute to visual perception, whereas those of Helmholtz were particularly relevant to top-down processes. Helmholtz's ideas about top-down expectations, or unconscious inferences, were precursors to the use of probabilistic Bayesian concepts by some recent vision theorists.

It is clear why probabilistic concepts are appealing in trying to understand mind and brain, if only because there is a lot of uncertainty in the world, and probability concepts try to conceptualize that uncertainty. But does the brain embody *classical* probability concepts, or does it deal with uncertainty in a new way? In this regard, it should be noted that the combination of shunting dynamics and competitive interactions with a RCF lead to automatic gain control and self-normalizing properties (Grossberg, 1973, 1980), so that the total activity of a processing channel in such a model is often approximately conserved. This total activity plays the role of a *real-time probability distribution*, that is, a deterministically computed distribution of normalized activity through time. Self-normalization enables such a network to maintain its sensitivity in response to distributed inputs whose total number and size can vary wildly through time. Grossberg (1973) called this property a solution of the *noise-saturation dilemma.* This property is useful when estimating the directional coherence of inputs whose total number and distribution can vary randomly through time, as occurs in the motion decision-making experiments. The robustness of this property when processing noisy data is illustrated by the fact that MODE model properties remain qualitatively unchanged when cellular noise is added to the model MT and MST processing stages. Indeed, all the processing stages of the MODE model—indeed all the models from the Grossberg school of modeling—use shunting neurons with self-normalizing properties in on-center off-surround networks. Thus self-normalizing neural dynamics capture key properties of classical probability theory, but they go beyond classical probability theory by being able to compute coherent, context-sensitive, time-varying measures that can exploit global properties of a scene.

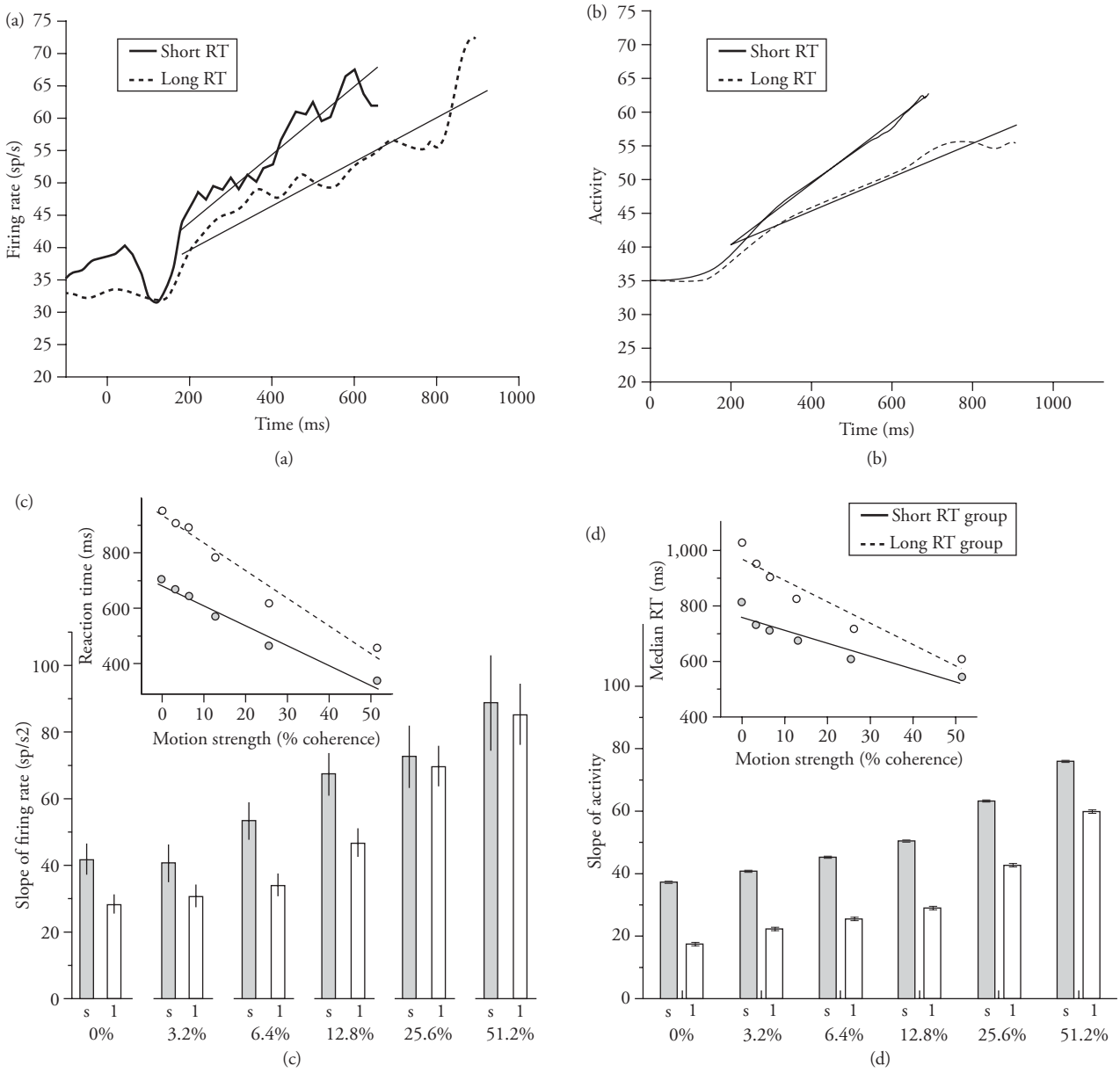


Figure I.7-22. Relationship between LIP response and reaction time. (a) Average LIP responses. (b) Model simulations reflect this relationship. (c) The histogram shows for each coherence level and group the slope of the linear fit to the average LIP response. (d) Model simulations reproduce these data trends. In panels (c) and (d), the error bars represent 95% confidence intervals. (Reprinted with permission from Grossberg & Pilly, 2008.)

When these self-normalizing "probabilistic" properties interact with the coherence-producing, aperture-problem-reducing, motion capture properties of the MT–MST grouping network, MODE model LIP cell activities covary quantitatively with the activation rate and amplitude of LIP cells that control a monkey's observable behavior in these experiments. The MODE model hereby clarifies how brain circuits that solve the aperture problem, notably the circuits that realize motion capture, control properties of probabilistic decision-making in real time. These directional decision-making data, and the motion capture process that enables it, can hereby be understood as a resolution of ambiguity that selects the best consensus movement that is compatible with motion data.

DOES BAYESIAN INFERENCE SHAPE BRAIN DESIGN?

The question of whether the mathematical laws of Bayesian inference shape brain design is of particular relevance to Newsome and Shadlen's data (e.g., Roitman & Shadlen, 2002; Shadlen & Newsome, 2001), since these investigators have proposed that perception and decision-making can be described as Bayesian inference, which

estimates the optimal interpretation of the stimulus given priors and likelihoods. Such an idea also seems appealing to people who believe that visual illusions are the brain's way of generating the optimal interpretation of the stimulus. However, Bayesian concepts do not provide a way to discover the brain mechanisms that make decisions, or indeed any behavioral and brain mechanisms. The Bayes rule, and related statistical concepts, are so general that they may be applied to problems in any science.

In particular, the Bayes rule follows from writing the probability of any two events I and S, namely p (I,S), in two different ways in terms of conditional probabilities, dividing both sides of this identity by a term on one side of the equation and then optimizing the result. In applications to perception, I is a stimulus that falls on the retina and the Bayesian model asks what percept S is most likely to have caused it. The basic idea is to find the S that maximizes the posterior probability p (S/I) via the Bayes rule.

This generality is part of its broad appeal but is also its weakness in not providing specific design principles or mechanisms that govern any particular science. Those who believe it can must ask themselves if the Bayes rule is sufficient, for example, to derive the laws of general relativity or high-energy physics. If not, then why can it do so for the brain?

The neural models that are discussed in this chapter have explained data that Bayesian models have heretofore failed to explain, do so in terms of basic brain design principles and mechanisms without an appeal to Bayesian inference, and, unlike other available models of these data, generate perceptual representations in response to the experimental visual stimuli. In particular, two neural design principles go into the MODE model, one that solves the aperture problem, thereby simulating neurophysiological data about interactions of cortical areas V1, MT, and MST, and the other that solves the noise-saturation dilemma and how this solution enables time-varying choices. Joining together these two designs enables simulation of the time course of LIP neurophysiological data, as well as of behavioral accuracy and reaction time properties, during both correct and error trials at different levels of input ambiguity in both fixed duration and reaction time tasks.

The succession of model processing stages V1–MT–MST embodies a hierarchical resolution of the directional uncertainties caused by the aperture problem. Model MST computes the global direction of random dot motion stimuli as the final stage of motion capture, while model LIP responds to MST outputs to compute the directional perceptual decision that leads to a saccadic eye movement. The intuitive idea is that the MT–MST feedback loop needs more time to capture the incoherent motion signals when there are more of them. In addition, this loop cannot achieve as high a level of asymptotic response magnitude, due to self-normalizing competition, when more incoherent motion signals compete with the emerging winning direction. In other words, the effectiveness of the motion capture process depends on input coherence. LIP then converts the inputs from MST into an eye movement directional command and thereby enables the (model) monkey to report its decision via a saccade. This self-organizing system thus trades accuracy against speed, with the brain running as fast as it can in response to the amount of uncertainty in the data that it is processing.

These various results cumulatively illustrate how cortical dynamics go beyond Bayesian concepts. They also clarify why probability theory ideas are initially so appealing in trying to explain visual percepts in general, and visual illusions in particular. Indeed, the brain is a self-organizing system that is designed to adapt autonomously to changing environments that include unexpected and rare events that may have no priors. These are not the types of processes for which Bayesian ideas are most appropriate. The brain needs to be studied and understood on its own terms, and we are well along the way to doing so, with visual illusions as one of many guides.

ART MODELS OF BRAIN LEARNING, RECOGNITION, AND CONSCIOUS PERCEPTION

If the Bayes rule is not the best way to represent top-down effects on perception, then what is? As the FACADE model was being developed, a parallel stream of modeling was aimed at understanding the neural mechanisms that support conscious experiences, notably how conscious visual percepts may arise. Such explanations are developed in Adaptive Resonance Theory (ART), which was introduced in Grossberg (1976a, 1976b). ART predicts that "all conscious states are resonant states." The 3D LAMINART theory describes a synthesis of ideas from the FACADE model about how 3D vision works in laminar visual cortical circuits that also embody ART mechanisms.

ART is a cognitive and neural theory of how the brain autonomously learns to attend, categorize, recognize, and predict objects and events in a changing world. To a remarkable degree, humans and other primates can rapidly learn new facts without being forced to just as rapidly forget what they already know. As a result, we can confidently go out into the world without fearing that, in learning to recognize a new friend's face, we will suddenly forget the faces of our family and friends. Otherwise expressed, ART avoids the problem of *catastrophic forgetting*.

ART accomplishes this by proposing a solution to the *stability-plasticity dilemma*, or how brains can learn quickly without also catastrophically forgetting already learned memories just as quickly. In so doing, ART clarifies key brain processes from which conscious experiences emerge. It describes a functional link between processes of consciousness, learning, expectation, attention, resonance, and synchrony (the CLEARS processes). ART predicted that all brain representations that solve the stability–plasticity dilemma use variations of CLEARS mechanisms (Grossberg, 1978, 1980, 2007, 2013, 2017). Through these CLEARS connections, ART clarifies why many animals are intentional beings who pay attention to salient objects, why "all conscious states are resonant states," and how brains can learn both *many-to-one maps* (representations whereby many object views, positions, and sizes all activate the same invariant object category)

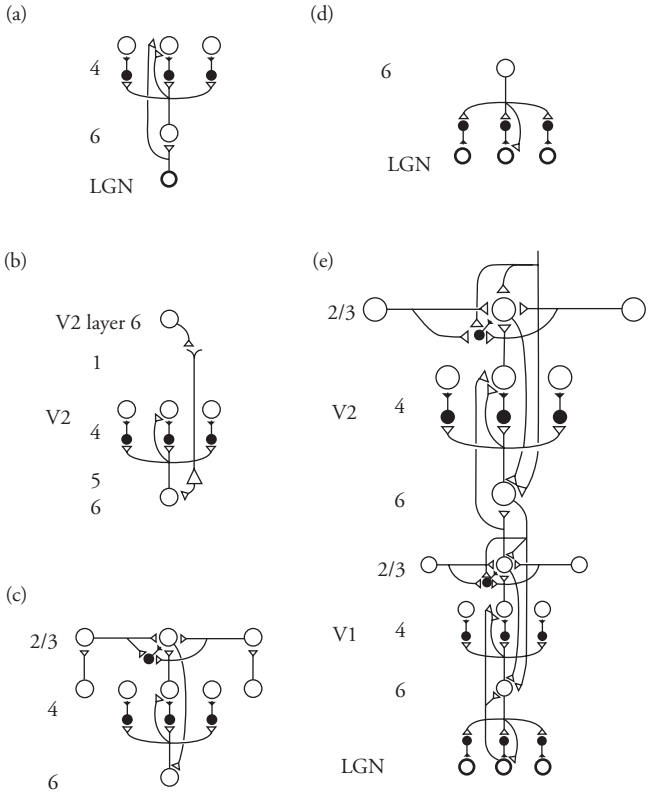


Figure I.7-23. Model LAMINART circuitry: How known cortical connections integrate bipole grouping with the ART Matching Rule for top-down attention. Inhibitory interneurons are shown filled-in black. (a) The LGN provides bottom-up activation to layer 4 via two routes. First, it makes a strong connection directly into layer 4. Second, LGN axons send collaterals into layer 6 and thereby also activate layer 4 via the 6 → 4 on-center off-surround path. This is an example of spatial competition. The combined effect of the bottom-up LGN pathways is to stimulate layer 4 via an on-center off-surround, which provides divisive contrast normalization. (b) "Folded feedback" carries attentional signals from higher cortex into layer 4 of V1, via the modulatory 6 → 4 path. Corticocortical feedback axons tend preferentially to originate in layer 6 of the higher area and to terminate in layer 1 of the lower cortex, where they can excite the apical dendrites of layer 5 pyramidal cells whose axons send collaterals into layer 6. The triangle in the figure represents such a layer 5 pyramidal cell. Several other routes through which feedback can pass into V1 layer 6 exist. Having arrived in layer 6, the feedback is then "folded" back up into the feedforward stream by passing through the 6 → 4 on-center off-surround path. This circuit realizes the top-down, modulatory on-center, off-surround circuit of the ART Matching Rule. (c) Connecting the 6 → 4 on-center off-surround to the layer 2/3 bipole grouping circuit: like-oriented layer 4 simple cells with opposite contrast polarities compete (not shown) before generating half-wave rectified outputs that converge onto layer 2/3 complex cells in the column above them. The long-range interactions within layer 2/3 realize the bipole grouping property. Just like attentional signals from higher cortex, as shown in (b), groupings that form among bipole cells in layer 2/3 also send activation into the folded feedback path, to enhance their own positions in layer 4 beneath them via the 6 → 4 on-center and to suppress input to other groupings via the 6 → 4 off-surround. There exist direct layer 2/3 → 6 connections in macaque V1, as well as indirect routes via layer 5. (d) Top-down corticogeniculate feedback from V1 layer 6 to LGN also has an on-center off-surround anatomy, similar to the 6 → 4 path, and realizes the ART Matching Rule from V1 to LGN. The on-center feedback selectively enhances LGN cells that are consistent with the activation that they cause, and the off-surround contributes to length-sensitive (endstopped) responses that facilitate grouping perpendicular to line ends. (e) The entire V1/V2 circuit: V2 repeats the laminar pattern of V1 circuitry but at a larger spatial scale. In particular, the horizontal layer 2/3 connections have a longer range in V2, allowing above-threshold perceptual groupings between more widely spaced inducing stimuli to form. V1 layer 2/3 projects up to V2 layers 6 and 4, just as LGN projects to layers 6 an 4 of V1. Higher cortical areas send feedback into V2, which ultimately reaches layer 6, just as V2 feedback acts on layer 6 of V1. Feedback paths from higher cortical areas straight into V1 (not shown) can complement and enhance feedback from V2 into V1. Top-down attention can also modulate layer 2/3 pyramidal cells directly by activating both the pyramidal cells and inhibitory interneurons in that layer. The inhibition tends to balance the excitation, leading to a modulatory effect. These top-down attentional pathways tend to synapse in layer 1. Their synapses on apical dendrites in layer 1 are not shown, for simplicity. (Reprinted with permission from Raizada & Grossberg, 2001.)

and *one-to-many maps* (representations that enable us to expertly know many things about individual objects and events).

ATTENTION, MATCHING, BIASED COMPETITION, RESONANCE, AND LEARNING

The views of Helmholtz enter the picture because ART accomplishes these properties by proposing how top-down expectations focus attention on salient combinations of cues, called "critical feature patterns", and characterizes how attention may operate via a form of self-normalizing "biased competition" (Desimone, 1998). In addition to anticipating the concept of biased competition by two decades, the exact form used for such attentive matching in explaining perceptual data (e.g., Bhatt, Carpenter, & Grossberg, 2007; Gove, Grossberg, & Mingolla, 1995) has been recently supported by the same form factor in the "normalization model of attention" (Reynolds & Heeger, 2009).

ART has developed Helmboltz's intuitions about unconscious inferences into rigorous science. Indeed, an active top-down expectation in ART plays the role of an unconscious inference. Top-down attentive matching in ART was predicted to be carried out by a top-down, modulatory on-center, off-surround network (Carpenter & Grossberg, 1987, 1991; Grossberg, 1980, 1999b) whose realization by identified cells within the laminar circuits of visual cortex has been described (Fig. I.7-23; Grossberg, 1999a; Grossberg & Versace, 2008; Raizada & Grossberg, 2001). The modulatory on-center in this top-down matching mechanism explains why unambiguous bottom-up data cannot usually be overridden by an active top-down expect at Matching Rule are combined with FACADE and 3D LAMINART mechanisms, many of the stresses between the views of Helmboltz and Kanizsa are resolved.

ART explains how such top-down attentive matching may help to solve the stability–plasticity dilemma by regulating cycles of *resonance* and *reset*, that is, of attentive matching and hypothesis testing (Fig. I.7-24). In particular, when a good enough match occurs between bottom-up

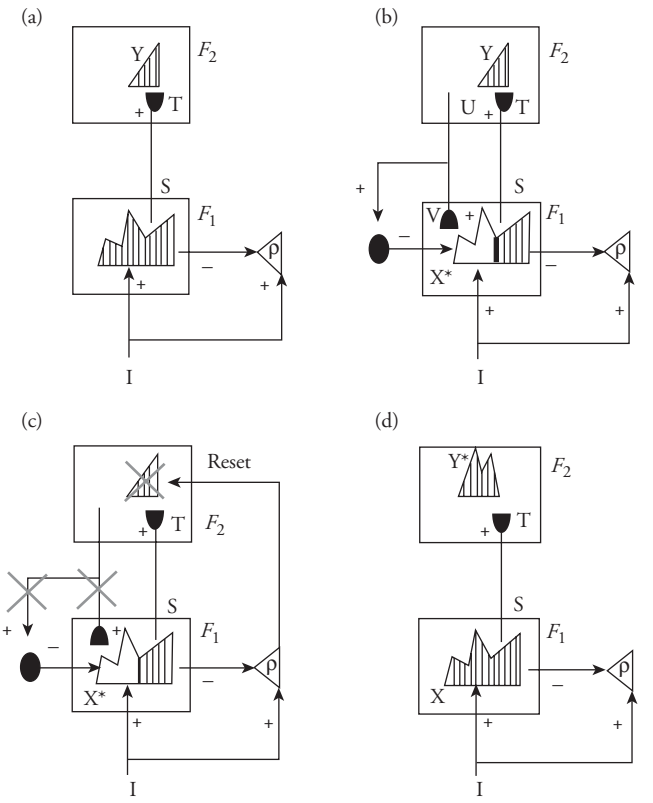


Figure I.7-24. How ART searches for a recognition category using cycles of resonance and reset. (a) Input pattern I is instated across feature detectors at level F_1 as an activity pattern X, while it nonspecifically activates the orienting system A with gain ρ. X inhibits A and generates output pattern S. S is multiplied by learned adaptive weights to form the input pattern T. T activates category cells Y at level F_2. (b) Y generates the top-down signals U, which are multiplied by adaptive weights and added at F_1 cells to form a *prototype* V that encodes the learned expectation of active F_2 categories. If V mismatches I at F_1, then a new STM activity pattern X^* (the hatched pattern) is selected at F_1. X^* is active at I features that are confirmed by V. Mismatched features (white area) are inhibited. When X changes to X^*, total inhibition decreases from F_1 to A. (c) If inhibition decreases sufficiently, A releases a nonspecific arousal burst to F_2; that is, "novel events are arousing." Arousal resets F_2 by inhibiting Y. (d) After Y is inhibited, X is reinstated and Y stays inhibited as X activates a different activity pattern Y^*. A search for better F_2 category continues until a better matching or novel category is selected. When the search ends, an attentive resonance triggers learning of the attended data. (Adapted with permission from Carpenter & Grossberg, 1993.)

inputs and a top-down expectation, then a synchronous resonant state emerges that embodies an attentional focus and is capable of driving fast learning of bottom-up recognition categories and top-down expectations—hence the name *adaptive* resonance. If the match is not good enough, then the currently active recognition category is reset by a complementary orienting system, and interactions between the attentional and orienting systems drive a search for a new, or better-matching, category, as schematized in Figure I.7-24.

All of the main ART predictions have received increasing support from psychological and neurobiological data since ART was introduced in 1976 (see Grossberg, 1999b, 2003; Grossberg & Versace, 2008; and Raizada & Grossberg, 2003, for reviews). Synchronous resonances are, in particular, expected to occur between multiple cortical and subcortical areas. Recent data support this prediction (e.g., see Buschman & Miller, 2007; Engel, Fries, & Singer, 2001; Grossberg, 2009b; and Pollen, 1999, for reviews).

CONSCIOUS VISUAL PERCEPTION IS A SURFACE-SHROUD RESONANCE

ART predicts that "all conscious states are resonant states," and FACADE theory predicts that all consciously visible percepts are surface percepts (Fig. I.7-6). This raises the question: What sort of resonances support conscious visible surface percepts, whether real or illusory? An answer to this question arose indirectly through a study of how the brain learns to recognize objects.

The ARTSCAN model clarifies how we can learn to recognize objects when we see them from multiple perspectives (e.g., Cao et al., 2011; Fazl et al., 2009). ARTSCAN shows how this may be accomplished by coordinating spatial attention and eye movements in the Where cortical processing stream (Goodale & Milner, 1992) with object attention and learning in the What cortical stream (Mishkin, Ungerleider, & Macko, 1983). A *surface-shroud resonance*, or synchronous feedback interaction, between cortical area V4 and parietal cortex can maintain spatial attention on an object's surface while the eyes explore it and drive the learning of an invariant object category. I have predicted that such a surface-shroud resonance supports conscious perception of the object's surface. After such a resonance locks attention upon a surface, it can also propagate top-down to lower visual areas, such as V1 and V2, and bottom-up to regions like prefrontal cortex, to synchronize them all to support the conscious percept.

HABITUATIVE GATES IN PERCEPTUAL AND RECOGNITION PROCESSES: AFTERIMAGES

Some visual illusions, such as afterimages, occur after the inducing stimuli shut off. Why do afterimages occur? In all the FACADE, 3D LAMINART, and ART models of perception and recognition, there exist activity-dependent habituative gating processes at the synapses of various connections. These gating processes may be realized either at presynaptic transmitters or postsynaptic receptors. The laws for habituative gates were introduced in Grossberg (1968, 1969). Neurobiological data and modeling in support of them from visual cortex were reported by Abbott, Varela, Sen, and Nelson (1997) in visual cortex and by Tsodyks and Markram (1997) in somatosensory cortex, using the names synaptic depression and dynamic synapses, respectively. These gating processes play several related functional roles. During the development of cortical maps, they prevent perseverative activation of the same cells and enable new inputs to learn how to activate new cells across the developing map

(e.g., Grossberg & Seitz, 2003; Olson & Grossberg, 1998). During percepts of changing visual inputs, they limit persistent activation of the same cells and enable percepts not to be pathologically smeared across a scene in response to moving objects (e.g., Francis & Grossberg, 1996a; Francis et al., 1994). During the learning of recognition categories, they enable the brain to reset categories whose top-down expectations mismatch bottom-up input patterns and to search for better-matching categories (e.g., Carpenter & Grossberg, 1987, 1991; Grossberg, 1980).

These adaptive properties also enable various illusions to occur. For example, the same activity-dependent habituative properties that limit persistence and help to drive the search for more predictive recognition categories can cause negative aftereffects and residual traces to be seen when images are suddenly shut off after being inspected for a while (Francis & Grossberg, 1996a). Habituative gating also plays a key role in generating bistable percepts, such as the Necker cube (Fig. I.7-16), since activity-dependent habituation of the pathways that support one percept can enable a competing percept to become dominant for a while (e.g., Grossberg & Swaminathan, 2004; Grossberg & Yazdanbakhsh, 2005; Grossberg et al., 2008).

In summary, both bottom-up and top-top mechanisms, and their intimate interaction, are needed to perceive and learn to recognize the world, and adaptive mechanisms of both types can cause the visual illusions that we recognize as such, or know, as well as those that we don't.

ACKNOWLEDGMENTS

This research was supported in part by the SyNAPSE program of DARPA (HR0011-09-C-0001).

REFERENCES

Abbott, L. F., Varela, K. Sen, K., & Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science*, *275*, 220–223.

Aubert, H. (1886). Die bewegungsempfinung. *Pfluggers Archive*, *39*, 347–370.

Baloch, A. A., & Grossberg, S. (1997). A neural model of high-level motion processing: Line motion and formotion dynamics. *Vision Research*, *37*, 3037–3059.

Beck, J., Prazdny, K., & Rosenfeld, A. (1983). A theory of textural segmentation. In J. Beck, B. Hope, & A. Rosenfeld (Eds.), *Human and machine vision* (pp. 1–38) New York: Academic Press.

Ben-Av, M. B., & Shiffrar, M. (1995). Disambiguating velocity estimates across image space. *Vision Research*, *35*, 2889–2895.

Berzhanskaya, J., Grossberg, S., & Mingolla, E. (2007). Laminar cortical dynamics of visual form and motion interactions during coherent object motion perception. *Spatial Vision*, *20*, 337–395.

Bhatt, R., Carpenter, G., & Grossberg, S. (2007). Texture segregation by visual cortex: Perceptual grouping, attention, and learning. *Vision Research*, *47*, 3173–3211.

Boardman, I., Grossberg, S., Myers, C., & Cohen, M. (1999). Neural dynamics of perceptual order and context effects for variable-rate speech syllables. *Perception & Psychophysics*, *6*, 1477–1500.

Born, R. T., & Tootell, R. B. H. (1992). Segregation of global and local motion processing in primate middle temporal visual area. *Nature*, *357*, 497–499.

Bowns, L. (2001). IOC, vector sum, and squaring: Three different motion effects or one? *Vision Research*, *41*, 965–972.

Brown, J., Bullock, D., & Grossberg, S. (1999). How the basal ganglia use parallel excitatory and inhibitory learning pathways to selectively respond to unexpected rewarding cues. *Journal of Neuroscience*, *19*, 10502–10511.

Brown, J. W., Bullock, D., & Grossberg, S. (2004). How laminar frontal cortex and basal ganglia circuits interact to control planned and reactive saccades. *Neural Networks*, *17*, 471–510.

Browning, A., Grossberg, S., & Mingolla, M. (2009a). A neural model of how the brain computes heading from optic flow in realistic scenes. *Cognitive Psychology*, *59*, 320–356.

Browning, A., Grossberg, S., & Mingolla, M. (2009b). Cortical dynamics of navigation and steering in natural scenes: Motion-based object segmentation, heading, and obstacle avoidance. *Neural Networks*, *22*, 1383–1398.

Buschman, T. J., & Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science*, *315*, 1860–1862.

Cao, Y., & Grossberg, S. (2005). A laminar cortical model of stereopsis and 3D surface perception: Closure and da Vinci stereopsis. *Spatial Vision*, *18*, 515–578.

Cao, Y., & Grossberg, S. (2012). Stereopsis and 3D surface perception by spiking neurons in laminar cortical circuits: A method of converting neural rate models into spiking models. *Neural Networks*, *26*, 75–98.

Cao, Y., Grossberg, S., & Markowitz, J. (2011). How does the brain rapidly learn and reorganize view-invariant and position-invariant object representations in inferior temporal cortex? *Neural Networks*, *24*, 1050–1061.

Carpenter, G. A., & Grossberg, S. (1987). A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, and Image Processing*, *37*, 54–115.

Carpenter, G. A., & Grossberg, S. (1993). Normal and amnesic learning, recognition, and memory by a neural model of cortico-hippocampal interactions. *Trends in Neurosciences*, *16*, 131–137.

Carpenter, G. A., & Grossberg, S. (Eds.). (1991). *Pattern recognition by self-organizing neural networks*. Cambridge, MA: MIT Press.

Castet, E., Lorenceau, J., Shiffrar, M., & Bonnet, C. (1993). Perceived speed of moving lines depends on orientation, length, speed and luminance. *Vision Research*, *33*, 1921–1936.

Chey, J., Grossberg, S., & Mingolla, E. (1997). Neural dynamics of motion grouping: From a perture ambiguity to object speed and direction. *Journal of the Optical Society of America A*, *14*, 2570–2594.

Chey, J., Grossberg, S., & Mingolla, E. (1998). Neural dynamics of motion processing and speed discrimination. *Vision Research*, *38*, 2769–2786.

Cisek, P. (2006). Integrated neural processes for defining potential actions and deciding between them: A computational model. *Journal of Neuroscience*, *26*, 9761–9770.

Cisek, P., Puskas, G. A., & El-Murr, S. (2009). Decisions in changing conditions: The urgency-gating model. *The Journal of Neuroscience*, *16*, 11560–11571.

Cohen, M. A., & Grossberg, S. (1984). Neural dynamics of brightness perception: Features, boundaries, diffusion, and resonance. *Perception & Psychophysics*, *36*, 428–456.

DeAngelis, G. C., Cumming, B. G., & Newsome, W. T. (1998). Cortical area MT and the perception of stereoscopic depth. *Nature*, *394*, 677–680.

Desimone, R. (1998). Visual attention mediated by biased competition in extrastriate visual cortex. *Philosophical Transactions of the Royal Society of London*, *353*, 1245–1255.

DeYoe, E. A., & Van Essen, D. C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neurosciences*, *11*, 219–226.

Elder, D., Grossberg, S., & Mingolla, E. (2009). A neural model of visually-guided steering, obstacle avoidance, and route selection. *Journal of Experimental Psychology: Human Perception & Performance*, *35*, 1501–1531.

Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top-down processing. *Nature Reviews Neuroscience*, *2*, 704–716.

Fang, L., & Grossberg, S. (2009). From stereogram to surface: How the brain sees the world in depth. *Spatial Vision*, *22*, 45–82.

Fazl, A., Grossberg, S., & Mingolla, E. (2009). View-invariant object category learning, recognition, and search: How spatial and object attention are coordinated using surface- based attentional shrouds. *Cognitive Psychology*, *58*, 1–48.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field." *Vision Research*, *33*, 173–193.

Filehne, W. (1922). Uber das optische Wahrnehmen von Bewegunggen. *Zeitschrift fur Sinnephysiologie*, *53*, 134–145.

Francis, G., & Grossberg, S. (1996a). Cortical dynamics of boundary segmentation and reset: Persistence, afterimages, and residual traces. *Perception*, *35*, 543–567.

Francis, G., & Grossberg, S. (1996b). Cortical dynamics of form and motion integration: Persistence, apparent motion, and illusory contours. *Vision Research*, *36*, 149–173.

Francis, G., Grossberg, S., & Mingolla, E. (1994). Cortical dynamics of feature binding and reset: Control of visual persistence. *Vision Research*, *34*, 1089–1104.

Goodale, M. A., & Milner, D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*, 10–25.

Gove, A., Grossberg, S., & Mingolla, E. (1995). Brightness perception, illusory contours, and corticogeniculate feedback. *Visual Neuroscience*, *12*, 1027–1052.

Grossberg, S. (1968). Some physiological and biochemical consequences of psychological postulates. *Proceedings of the National Academy of Sciences USA*, *60*, 758–765.

Grossberg, S. (1969). On the production and release of chemical transmitters and related topics in cellular control. *Journal of Theoretical Biology*, *22*, 325–364.

Grossberg, S. (1973). Contour enhancement, short term memory, and constancies in reverberating neural networks. *Studies in Applied Mathematics*, *52*, 217–257.

Grossberg, S. (1976a). Adaptive pattern classification and universal recoding: I. Parallel development and coding of neural feature detectors. *Biological Cybernetics*, *23*, 121–134.

Grossberg, S. (1976b). Adaptive pattern classification and universal recoding: II. Feedback, expectation, olfaction, and illusions. *Biological Cybernetics*, *23*, 187–202.

Grossberg, S. (1978). A theory of human memory: Self-organization and performance of sensory-motor codes, maps, and plans. In R. Rosen & F. Snell (Eds.), *Progress in theoretical biology* (Vol. 5, pp. 233–374). New York: Academic Press.

Grossberg, S. (1980). How does a brain build a cognitive code? *Psychological Review*, *87*, 1–51.

Grossberg, S. (1984). Outline of a theory of brightness, color, and form perception. In E. Degreef & J. van Buggenhaut (Eds.), *Trends in mathematical psychology*. Amsterdam: North-Holland.

Grossberg, S. (1987). Cortical dynamics of three-dimensional form, color, and brightness perception: II. Binocular theory. *Perception & Psychophysics*, *41*, 117–158.

Grossberg, S. (1991). Why do parallel cortical systems exist for the perception of static form and moving form? *Perception & Psychophysics*, *49*, 117–141.

Grossberg, S. (1994). 3-D vision and figure-ground separation by visual cortex. *Perception & Psychophysics*, *55*, 48–120.

Grossberg, S. (1997). Cortical dynamics of three-dimensional figure–ground perception of two-dimensional figures. *Psychological Review*, *104*, 618–658.

Grossberg, S. (1998). How is a moving target continuously tracked behind occluding cover? In T. Watanabe (Ed.), *High level motion processing: Computational, neurobiological, and psychophysical perspectives* (pp. 3–52). Cambridge, MA: MIT Press.

Grossberg, S. (1999a). How does the cerebral cortex work? Learning, attention, and grouping by the laminar circuits of visual cortex. *Spatial Vision*, *12*, 163–187.

Grossberg, S. (1999b). The link between brain learning, attention, and consciousness. *Consciousness and Cognition*, *8*, 1–44.

Grossberg, S. (2000). The complementary brain: Unifying brain dynamics and modularity. *Trends in Cognitive Sciences*, *4*, 233–246.

Grossberg, S. (2003). How does the cerebral cortex work? Development, learning, attention, and 3D vision by laminar circuits of visual cortex. *Behavioral and Cognitive Neuroscience Reviews*, *2*, 47–76.

Grossberg, S. (2007). Consciousness CLEARS the mind. *Neural Networks*, *20*, 1040–1053.

Grossberg, S. (2008). The art of seeing and painting. *Spatial Vision*, *21*, 463–486.

Grossberg, S. (2013). Adaptive Resonance Theory: How a brain learns to consciously attend, learn, and recognize a changing world. *Neural Networks*, *37*, 1–47.

Grossberg, S. (2014). How visual illusions illuminate complementary brain processes: Illusory depth from brightness and apparent motion of illusory contours. *Human Neuroscience,* doi:10.3389/fnhum.2014.00854/full.

Grossberg, S. (2017). Towards solving the Hard Problem of Consciousness: The varieties of brain resonances and the conscious experiences that they support. *Neural Networks*, *87*, 38–95. http://www.sciencedirect.com/science/article/pii/S0893608016301800

Grossberg, S., & Howe, P. D. L. (2003). A laminar cortical model of stereopsis and three- dimensional surface perception. *Vision Research*, *43*, 801–829.

Grossberg, S., Markowitz, J., & Cao, Y. (2011). On the road to invariant recognition: Explaining tradeoff and morph properties of cells in inferotemporal cortex using multiple-scale task-sensitive attentive learning. *Neural Networks*, *24*, 1036–1049.

Grossberg, S., & McLoughlin, N. (1997). Cortical dynamics of three-dimensional surface perception: Binocular and half-occluded scenic images. *Neural Networks*, *10*, 1583–1605.

Grossberg, S., & Mingolla, E. (1985a). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review*, *92*, 173–211.

Grossberg, S., & Mingolla, E. (1985b). Neural dynamics of perceptual grouping: Textures, boundaries and emergent segmentations. *Perception & Psychophysics*, *38*, 141–171.

Grossberg, S., & Mingolla, E. (1987). Neural dynamics of surface perception: Boundary webs, illuminants, and shape-from-shading. *Computer Vision, Graphics and Image Processing*, *37*, 116–165.

Grossberg, S., Mingolla, E., & Pack, C. (1999). A neural model of motion processing and visual navigation by cortical area MST. *Cerebral Cortex*, *9*, 878–895.

Grossberg, S., Mingolla, E., & Viswanathan, L. (2001). Neural dynamics of motion integration and segmentation within and across apertures. *Vision Research*, *41*, 2521–2553.

Grossberg, S., & Pessoa, L. (1998). Texture segregation, surface representation, and figure-ground separation. *Vision Research*, *38*, 2657–2684.

Grossberg, S., & Pilly, P. (2008). Temporal dynamics of decision-making during motion perception in the visual cortex. *Vision Research*, *48*, 1345–1373.

Grossberg, S., & Pinna, B. (2012). Neural dynamics of Gestalt principles of perceptual organization: From grouping to shape and meaning. *Gestalt Theory*, *34*, 399–482.

Grossberg, S., Roberts, K., Aguilar, M., & Bullock, D. (1997). A neural model of multimodal adaptive saccadic eye movement control by superior colliculus. *Journal of Neuroscience*, *17*, 9706–9725.

Grossberg, S., & Rudd, M. E. (1992). Cortical dynamics of visual motion perception: Short-range and long-range apparent motion. *Psychological Review*, *99*, 78–121.

Grossberg, S., & Seitz, A. (2003). Laminar development of receptive fields, maps, and columns in visual cortex: The coordinating role of the subplate. *Cerebral Cortex*, *13*, 852–863.

Grossberg, S., & Swaminathan, G. (2004). A laminar cortical model for 3D perception of slanted and curved surfaces and of 2d images: Development, attention, and bistability. *Vision Research*, *44*, 1147–1187.

Grossberg, S., & Todorović, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena. *Perception & Psychophysics*, *43*, 241–277.

Grossberg, S., & Versace, M. (2008). Spikes, synchrony, and attentive learning by thalamocortical circuits. *Brain Research*, *1218*, 278–312.

Grossberg, S., & Yazdanbakhsh, A. (2005). Laminar cortical dynamics of 3D surface perception: Stratification, transparency, and neon color spreading. *Vision Research*, *45*, 1725–1743.

Grossberg, S., Yazdanbakhsh, A., Cao, Y., & Swaminathan, G. (2008). How does binocular rivalry emerge from cortical mechanisms of 3-D vision? *Vision Research*, *48*, 2232–2250.

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, *23*, 141–221.

Kelly, F., & Grossberg, S. (2000). Neural dynamics of 3-D surface perception: Figure–ground separation and lightness perception. *Perception & Psychophysics*, *62*, 1596–1618.

Lisberger, S. G., Morris, E. J., & Tychsen, L. (1987). Visual motion processing and sensory-motor integration for smooth pursuit eye movements. *Annual Review of Neuroscience*, *10*, 97–129.

Livingstone, M. S., & Hubel, S. H. (1984). Anatomy and physiology of a color system in the primate visual cortex. *Journal of Neuroscience*, *4*, 309–356.

Livingstone, M. S., & Hubel, S. H. (1988). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science*, *240*, 740–749.

Lorenceau, J., & Gorea, A. (1989). "Blobs" are critical in perceiving the direction of moving plaids. *Perception*, *18*, 539.

Lorenceau, J., & Shiffrar, M. (1992). The influence of terminators on motion integration across space. *Vision Research*, *32*, 263–273.

Lueschow, A., Miller, E. K., & Desmone, R. (1994). Inferior temporal mechanisms for invariant object recognition. *Cerebral Cortex*, *4*, 523–531.

Marr, D. & Ullman, S. (1981). Directional selectivity and its use in early visual processing. *Proceedings of the Royal Society of London, B*, *211*, 151–180.

Matisse, H. (1992). *Jazz*. Scranton, PA: George Braziller. (Original work published 1947)

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences*, *6*, 414–417.

Olson, S., & Grossberg, S. (1998). A neural network model for the development of simple and complex cell receptive fields within cortical maps of orientation and ocular dominance. *Neural Networks*, *11*, 189–208.

Pack, C. C., & Born, R. T. (2001). Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. *Nature*, *409*, 1040–1042.

Pack, C. C., Grossberg, S., & Mingolla, E. (2001). A neural model of smooth pursuit control and motion perception by cortical area MST. *Journal of Cognitive Neuroscience*, *13*, 102–120.

Peterhans, E., & von der Heydt, R. (1989). Mechanisms of contour perception in monkey visual cortex: II. Contours bridging gaps. *Journal of Neuroscience*, *9*, 1749–1763.

Pinna, B. (1987). Un effetto di colorazione. In V. Majer, M. Maeran, & M. Santinello (Eds.), *Il laboratorio e la città: XXI Congresso degli Psicologi Italiani* (p. 158). Milano: Societá Italiana di Psiocologia, 1987.

Pinna, B. (2009). New Gestalt principles of perceptual organization: An extension from grouping to shape and meaning. *Gestalt Theory*, *32*, 11–78.

Pinna, B., Brelstaff, G., & Spillmann, L. (2001). Surface color from boundaries: A new "watercolor" illusion. *Vision Research*, *41*, 2669–2676.

Pinna, B., & Grossberg, S. (2005). The watercolor illusion and neon color spreading: A unified analysis of new cases and neural mechanisms. *Journal of the Optical Society of America A*, *22*, 2207–2221.

Pollen, D. A. (1999). On the neural correlates of visual perception. *Cerebral Cortex*, *9*, 4–19.

Ponce, C. R., Lomber, S. G., & Born, R. T. (2008). Integrating motion and depth via parallel pathways. *Nature Neuroscience*, *11*, 216–223.

Raizada, R. D. S., & Grossberg, S. (2001). Context-sensitive bindings by the laminar circuits of V1 and V2: A unified model of perceptual grouping, attention, and orientation contrast. *Visual Cognition*, *8*, 431–466.

Raizada, R. D. S., & Grossberg, S. (2003). Towards a theory of the laminar architecture of cerebral cortex: Computational clues from the visual system. *Cerebral Cortex*, *13*, 100–113.

Ramachandran, V. S. (1985). Apparent motion of subjective surfaces. *Perception*, *14*, 127–134.

Ramachandran, V. S., Rao, V. M., & Vidyasagar, T. R. (1973). Apparent motion with subjective contours. *Vision Research*, *13*, 1398–1401.

Reynolds, J. H., & Heeger, D. J. (2009). The normalization model of attention. *Neuron*, *61*, 168–185.

Roitman, J. D., & Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of Neuroscience*, *22*, 9475–9489.

Shadlen, M. N., & Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *Journal of Neurophysiology*, *86*, 1916–1936.

Shimojo, S., Silverman, G. H., & Nakayama, K. (1989). Occlusion and the solution to the aperture problem for motion. *Vision Research*, *29*, 619–26.

Tse, P. U. (1999). Volume completion. *Cognitive Psychology*, *39*, 37–68.

Tsodyks, M. V., & Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proceedings of the National Academy of Sciences USA*, *94*, 719–723.

Usher, M., & McCllelland, J. L. (2001). The time course of perceptual choice: the leaky competing accumulator model. *Psychological Review*, *108*, 550–592.

Van Tuijl, H. F. J. M. (1975). A new visual illusion: Neonlike color spreading and complementary color induction between subjective contours. *Acta Psychologica*, *39*, 441–445.

Varin, D. (1971). Fenomini di contrasto e diffusione chromatica nell organizzazone spaziale del campo percettivo. *Revista di Psychologica*, *65*, 101–128.

von der Heydt., R., Peterhans, E., & Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science*, *224*, 1260–1262.

von Holst, E. (1954). Relations between the central nervous system and the peripheral organs. *British Journal of Animal Behaviour*, *2*, 89–94.

Wallach, H. (1935). On the visually perceived direction of motion. *Psychologische Forschung*, *20*, 325–380.

Wuerger, S., Shapley, R., & Rubin, N. (1996). "On the visually perceived direction of motion" by Hans Wallach: 60 years later. *Perception*, *25*, 1317–1367.

Young, L. R., Forster, J. D., & van Houtte, N. (1968). A revised stochastic sampled model for eye tracking movements. Paper presented at the 4th Annual NASA-University Conference on Manual Control, University of Michigan, Ann Arbor.

Zoccolan, D., Kouh, M., Poggioa, T., & DiCarlo, J. J. (2007). Trade-off between object selectivity and tolerance in monkey inferotemporal cortex. *Journal of Neuroscience*, *27*, 12292–12307.

Chapter 8
Visual Illusions?

Jan Koenderink

ILLUSIONS?

Illusion is commonly defined as something that deceives by producing a false or misleading impression of reality. In psychology it is a perception that represents what is perceived in a way different from the way it is in reality. Illusion differs from *delusion*, which is commonly defined as a false belief or opinion. In psychiatry it is a fixed false belief that is resistant to reason or confrontation with actual fact. Illusion also differs from *deception*, which is defined as something that is intended to deceive, like fraud, or artifice. Notice that illusions necessarily refer to immediate visual awareness, whereas delusions and deceptions (like fantasies, hallucinations, or psychotic beliefs) need not; they can just as well occur in the cognitive domain.

These common notions set the focus for this chapter. Illusion, delusion, and deception all allude to erroneous mental representations. Instances of the latter two are often denoted "illusions," but this is unfortunate. The ontological differences are essential.

A surprising conjurer's trick (Kuhn, Amiani, & Rensink, 2008) is a *deception*. Deceptions always result from intentions of other people, like the magician. The magician intentionally tries to deceive the audience. In order to be successful at this, the "illusionist" applies a practical understanding of visual perception. Deceptions are different from illusions in that they go away once one "knows the trick." Thus, in a deeper sense, they have nothing to do with perception proper. They are best considered as lies, or fiction. Deceptions are not considered any further here.

Seeing blue mice and pink elephants in *delerium tremens* (London, 1913/1989)[1] is a case of *delusion*. Such visions are personal—others do not necessarily experience them. For this reason they are also known as *hallucinations*. If the majority experiences a flying saucer it counts as an observation, whereas if a single person defends such an experience against the majority opinion it counts as a hallucination.[2] Hallucinations commonly occur in perception, as when one takes a rope for a snake (Fig. I.8-1). Immediate awareness is precategorical and may have contained a kind of rope-snake-like entity.[3] The perception involves a categorical judgment and is in the cognitive domain. Because they are reflective thought rather than truly perceptual, delusions are also not considered any further here.

Illusions occur in immediate visual experience following immediate visual awareness. Awareness may be called "presentation" so as to differentiate it from "representation." Presentations simply *happen to a person*; they are protocognitive. They differ from thoughts, which are *actions of a person*. Presentations as such cannot be illusionary, because they are beyond true or false. The perception, or visual experience, that immediately follows the presentation is intentional in the sense that it refers to physical reality. Consequently, it may turn out to be nonveridical, in which case one speaks of an "optical illusion."

Illusions in the sense intended in this chapter are not under voluntary control. They stubbornly persist in immediate awareness, despite the fact that the conscious perception is evidently illusory. For instance, in the Zöllner (1860) illusion, lines that are geometrically parallel on the paper (perhaps one has drawn them so oneself) stubbornly appear to be converging (Fig. I.8-2). This is a major philosophical reason to be interested in such cases. They appear to threaten the very nature of our reality.

Notice that the notion of illusion implies the existence of an objective reality, a world that exists independent of the senses, and one that is not dependent on (and thus the same for any) observer. It is a (metaphysical) objectivist and projectivist view. Such a view is described by David Hume (1777/1975, I.iii.XIV):

> 'Tis a common observation, that the mind has a great propensity to spread itself on external objects, and to conjoin with them any internal impressions, which they occasion, and which always make their appearance at the same time that these objects discover themselves to the senses.

(See Fig. I.8-3.) There are severe problems with such a view. It is not coherent with modern physics, and it is overly anthropocentric. Moreover, it has led to numerous unfortunate misconceptions in psychology (Lakoff, 1987). However, for the purposes of this chapter, these (very real and serious) problems are not directly addressed.

An "optical illusion" may be understood as an immediate visual experience whose structure differs from the structure of the environment as monitored by other means, such as touch, audition, taste, or smell, augmented with pointer readings from a variety of physical instruments. The major



Figure I.8-1. Taking a rope for a snake, an ancient Vedanta (Deussen, 1883) metaphor. (http://sahajapower.files.wordpress.com/2008/04/rope-snake.jpg)



Figure I.8-2. The weird effect seen in this building, a part of Melbourne's Digital Harbour Port, depends on a variety of the Zöllner illusion often known as the "Café wall illusion" (Gregory & Heard, 1979). (https://stampsy.com/stamp/20319)

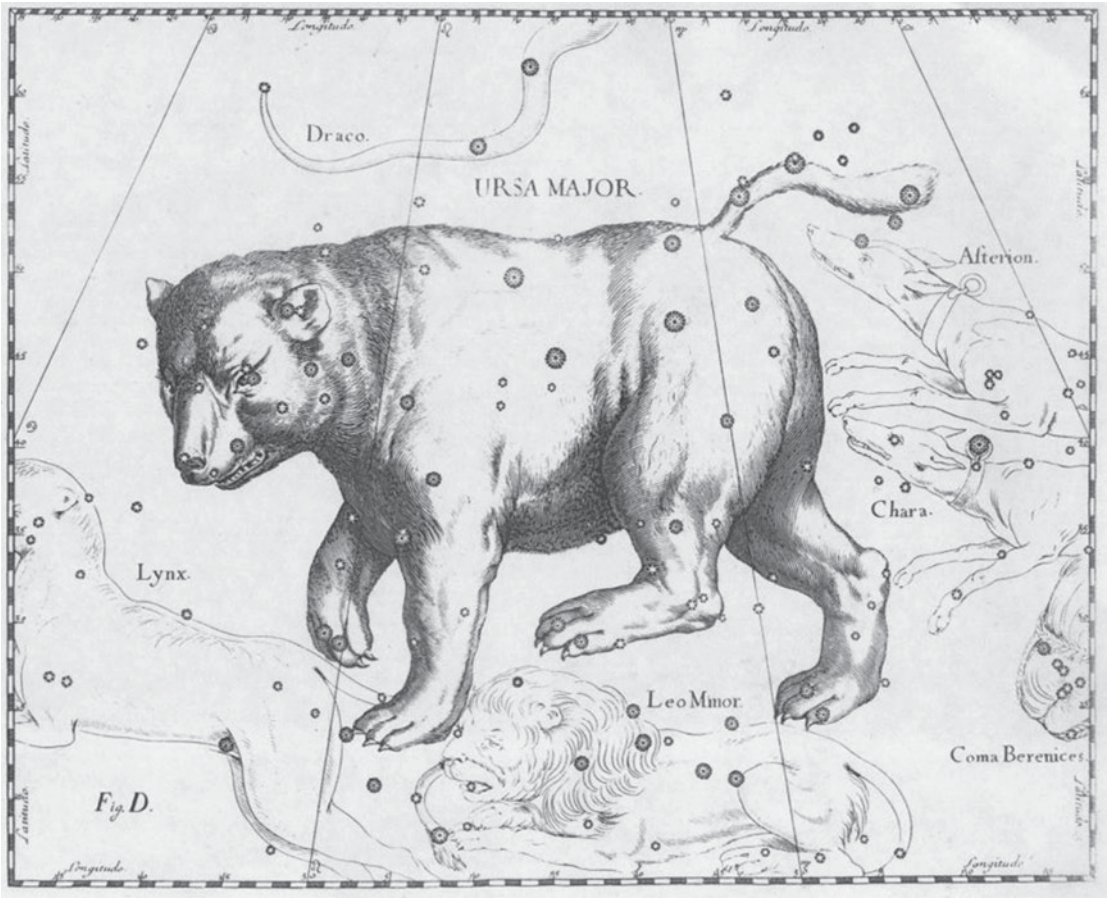


Figure I.8-3. The constellation of Ursa Major (the Big Dipper) drawn by Johannes Hevelius (1690). (Notice that the drawing is as being viewed from the outside of the celestial sphere.) This illustrates David Hume (1777/1975): "the mind has a great propensity to spread itself on external objects." Are stellar constellations illusions? Something to ponder.

problem with this approach is that it is not obvious which perceptual structures are supposed to be correlated with which physical structures. This is because the qualia of awareness and the physical properties of stimuli exist on ontologically different levels of reality (Poli, 2006). For instance, it is not evident that color (a quale; i.e., a mental entity) would correlate with either the radiant power spectrum (an electromagnetic entity) or the spectral reflectance (a material property). In practice it does not correlate very well with either. This is problematic from the projectivist perspective.

Similar observations apply to virtually all qualia, whether "primary" (shapes, sizes, etc.) or "secondary" (colors, etc.). It is the reason for John Locke's (1690) notion of primary and secondary qualities:

> Such qualities, which in truth are nothing in the objects themselves, but powers to produce various sensations in us by their primary qualities, i.e. by the bulk, figure, texture, and motion of their insensible parts as colours, sounds, tasts, etc. These I call secondary qualities. (II.VIII.§10)

(See Fig. I.8-4.) However, from the perspective of the phenomenology of visual awareness, Locke's "primary" qualities are on the same ontological level as his "secondary" ones; this is evident from the numerous geometrical illusions. It was already clear to Bishop Berkeley (1709). The dichotomy fails to make sense in view of the phenomenological facts.

As is often related in popular books on vision, a piece of coal in sunlight sends more radiation to the eye than a piece of white paper in shadow (Metzger, 1936). Some authors construe this as an illusion (we see wildly nonveridical radiances) and others as a sign that our vision is not fooled by incidental circumstances (light and shadow) and that we see the coal and paper "as they really are." Obviously both viewpoints have something going for them. If we assume a strong stand on the issue, we apparently are wondering "how God *should* have set up things": it is again based on the objectivist and projectivist stance. Examples of cases like this multiplied after computer vision provided us with the numerous "Shape from X" theories (Forsyth & Ponce, 2003). Showing that human observers fail to function according to these algorithms, at least in detail, has grown to something like a cottage industry in the human vision community. It is a rich mine for novel "illusions" if one is so minded.

Evidently then, the topic of "optical illusion" is a very hairy one. It depends on one's ontological convictions. Some would say the primordial optical illusion is the scene that magically appears in awareness when we open our eyes in broad daylight. One way to live with this is to equate awareness and reality. Others are convinced that science has told us what reality is and that science will eventually also tell us how it comes about that the presentations in awareness are such close representations of the world as it actually is. These convictions are mutually exclusive, but they determine what illusions are for

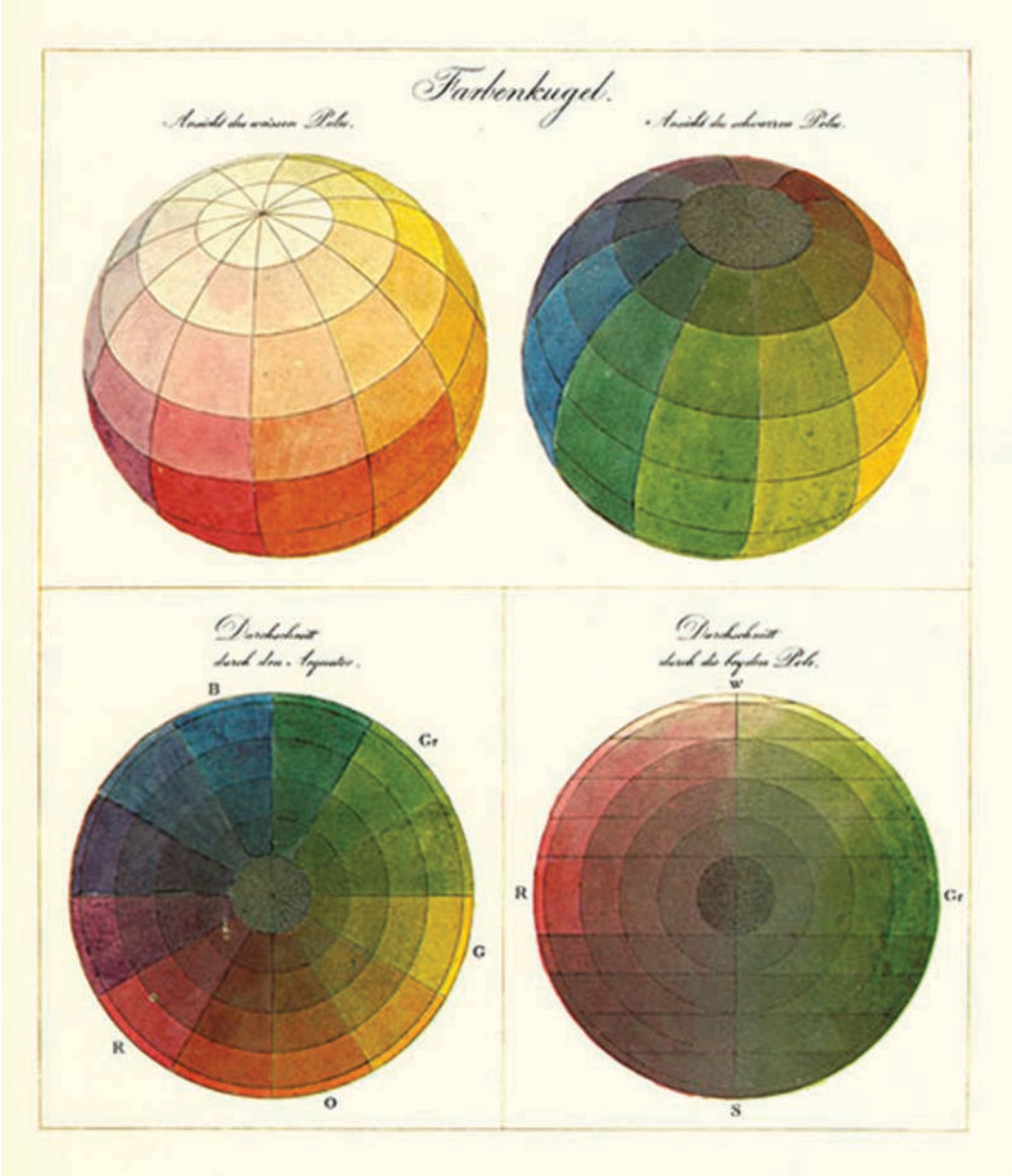


Figure I.8-4. The "color sphere" as drawn by Philip Otto Runge (1810), a Danish painter of the early 19th century. Runge was (other than being a painter) a mystic, who understood colors in terms of their Christian significance. Although the first example of a scientifically correct (at least in a topological sense) color body, these colors are very much secondary qualities: "such are nothing in the objects themselves, but powers to produce various sensations in us." Are colors illusions, delusions, or hallucinations, granted that they are nothing physical? Are they mere mental paint and correspond to nothing in our environment?

each of us and what illusions can teach us about vision. In order to cater to such widely divergent convictions, this chapter briefly considers alternatives to overall approaches to vision and attempts a taxonomy of various types of illusions.

MICROGENY

"Vision" is such a complex issue that it is virtually impossible to list all distinct approaches to it. For the sake of this short chapter on illusions, we use a caricatural dichotomy involving only the extreme views. These extremes are perhaps "inverse optics" (Marr, 1982; Poggio, 1990) and "controlled hallucination" (Hoffman, 2009; Koenderink, 2011). Both limits are (intentionally) caricatures. Most vision scientists would subscribe to a position somewhere in between, with a minority found at either extreme. Each position implies very different stances on many issues that are of key importance for the topic of illusions. This is the problem of the microgeny of visual awareness (Brown, 1977; Koenderink, 2011).

Inverse optics captures the idea that the radiances at the corneas of the eyes contain all information (often denoted "optical data") to compute a veridical representation of the physical scene in front of the eyes. All that is needed is sufficient computational power. Thus the visual system consists of a bunch of algorithms. A recent refinement is

to relax determinism and let the algorithms be Bayesian estimators (Knill & Richards, 1996). This allows prior experience to be used (as Bayesian priors) and ensures the best possible estimates. In any case, the microgeny of presentations involves mere symbol pushing. The meaning and qualities in awareness cannot be explained; here one takes one of three obvious solutions: —the *first* and perhaps most consistent solution is to simply deny the very existence of meanings and qualities (behaviorism); —the *second*, perhaps "mystical" solution is to hold that the meanings and qualities are already present in the optical data (hence "data" instead of "structure"), so there is no need to compute them; —the *third*, perhaps "magical" solution is to hold that sufficient computation is certain to lift mere structure to the level of meaning and quality. The second solution is *mystical* because it is fully incompatible with the sciences. The third solution is *magical* because it needs some kind of *deus ex machina* mechanism ("emergent properties," "center of consciousness," or some mysterious quantum events [Penrose, 1989/1999], and so forth). The first solution is fine as far as it goes, which is unfortunately not very far. If one is satisfied that (scientifically speaking) speech is just a movement of air molecules (Watson, 1930), the matter is settled. Perhaps surprisingly, many people find this a congenial notion. If so, we are done here.

The *inverse optics* microgeny is often denoted as "bottom-up." This view may perhaps be characterized as the "Marrian approach" (Marr, 1982). It results in *representations* of the scene in front of the observer. There is no natural distinction between a real-time system that runs beyond awareness and a meaning-creating system that results in conscious experiences. The approach is consistent with the notion that consciousness is a spooky thing that will soon be replaced with some scientific notion.

Controlled hallucination is often denoted as "top-down." Microgeny proceeds by the progressive diversification and pruning of many diverse threads of proto-awareness, a kind of evolutionary process (Brown, 1977). The "pruning" proceeds through checks against the optical structure stored in volatile front-end buffers, a kind of top-down probing. Because of this pruning the finally selected awareness will be a good "explanation" of that optical structure. Controlled hallucination is an information-creating process (Koenderink, 2011), thus the problems with meaning and qualia are irrelevant. The point is that the pruning is implemented as a probing of the front-end structure (and thus, by long chains of indirection, of the physical structure of the environment), essentially a kind of "questioning" (Schrödinger, 1943). Questions imply possible answers, and the meaning gained by the answer is defined by the question, rather than the answer. One way to understand this is to regard the question (or probing) as a "format" in the sense of computer science.[4] Another way to understand it is to regard it as an analogue of the scientific method.

Such probabilistic "analysis-by-synthesis" methods (Yuille & Kersten, 2006) are the only information-creating processes known (Koenderink, 2011). Deterministic algorithms merely transform structure into structure and are forever unable to cross the border between structure and meaning. Controlled hallucination fits current biological thought (ethology) seamlessly. In psychology, such methods were recently (re)introduced as a "user interface model" of perception by Hoffman (2009) and Koenderink (2011) and are not that different from von Uexküll's (1909) notion of *Umwelt*. This view may perhaps be characterized as the ethological approach (Lorenz, 1973; Riedl, 1980; Tinbergen, 1951). It results in *presentations* that aid the agent in its *potential* actions, the system being complementary to a *real-time* system that runs outside of awareness. The approach accepts awareness as an aspect of reality that cannot be explained away on the basis of present or future physics. The latter notion is of course not palatable to tough-minded scientists.

In both views, the nature of vision is such that it comes in *glimpses* that are generated at a timescale of less than a tenth of a second, whereas a *glance* comprises maybe a dozen glimpses. A glance thus occurs on the timescale of about a second and may involve a number of distinct involuntary fixations. In such a "specious present," the temporal order of glimpses is not necessarily retained. A *good look* comprises maybe a dozen glances, thus lives at a timescale of up to 10 s. *Scrutinizing* involves any number of good looks. In a good look, the glances are retained in their temporal order. What is in a glimpse depends on the momentary fixation. This is important, because sampling in the periphery is sparse and necessarily lossy. The glimpses and glances are somehow integrated in a perception that seems complete and of uniformly high resolution to the observer. It is a most compelling delusion. Most of the fixations are involuntary, though—especially in scrutinizing—the observer probes through voluntary fixations. The nature of the process that leads to an integrated perception on the basis of glimpses and glances is still very ill understood. Although both views of the nature of vision interpret these facts in different ways, both necessarily deal with them. One aspect of the integration that appears more or less settled is that the spatiotemporal integration is based on relative, local structures rather than a global, overall framework. It has come as a revelation to the recent mainstream that there simply *is no* global image buffer at foveal resolution (Rensink, 2002, 2004; Simons & Rensink, 2005).

Notice that numerous "in between" models are possible. The controlled hallucination needs to be booted up so to speak, so it is perhaps natural to think of an independent "gist system" that would run in bottom-up fashion (Oliva & Torralba, 2006). However, it is equally possible to think of parallel processes at various levels of "detailedness" (more general than "resolution") but each essentially top-down. Since there are only a limited number of "gists," the hallucinations might as well start at random and still be effective. This may be likened to the game of "twenty questions."[5] (The initial question is necessarily a random one, often a fixed one like "animal, vegetable, mineral?") Moreover, in virtually all cases the setting changes little from moment to moment, so the gist is settled beforehand anyway. In the inverse optics view, a "gist" system would serve as a top-down control. One might also conceive of bottom-up constructions of "proto-objects" (Clark, 2004), also in the controlled hallucination view. However, in that view these would be meaningless data-structures, merely serving to speed up "constraint seeking," that is, the probing of the

microgenetic process. In the inverse optics view, the proto-objects would be an initial stage in an overall bottom-up process. There is hardly a bound on the variety of *in between* models that might be proposed.

Perhaps the most essential distinction (with regard to the topic of illusions) is the existence of a representation of the scene in front of the observer. In the conventional view the correspondence between the representation and the physical scene determines whether vision is veridical or illusory. There are many problems here: it is not at all clear whether any such representations exist, and it depends on what one means by "exist," "representation," and "physical world." It has proven fruitless to try to reach consensus on the meaning of these terms. Although the real-time, subconscious system is often thought to maintain a largely veridical representation, as evident from numerous successful visuomotor actions, the conclusion is by no means necessary. The system may well be close to the subsumption architectures proposed by Rodney Brooks (1991) for robotics, for instance.

The holy grail of inverse optics is indeed a veridical representation, but such a representation would necessarily be without meanings and qualities. In order to make things stick, one would have to describe the illusions also in purely objective terms. This is sometimes possible, although many illusions are outside the range of such an approach. Controlled hallucination yields presentations that are by their very nature meaningful and composed of qualities. However, the notion of a God's eye view does not fit this model. Thus either everything is illusory or nothing is—a mere matter of definition. Here one might perhaps say that the "user interface" has greater or lesser *utility*, so one has *degrees of illusiveness*.[6] However, utility is hard to assess in the presence of a real-time, subconscious system that takes care of much of the visuomotor action. Thus the topic of "representation" remains a muddled one.

In biology, more specifically ethology, one finds numerous descriptions of instances of—from a human perspective—*strange*, but for the animal's behaviour, very *effective* user-interfaces (Lorenz, 1973; Tinbergen, 1951). For example, the common cattle tick takes the conjunction of warmth and the smell of butyric acid for a mammal, the lapwing goose takes a brick, located near the nest, for an egg, the female stickleback fish, in the breeding season, takes anything red for a male (Fig. I.8-5), and so forth. At least, judging from their actions. Are these animals deluded, or suffering from illusions? In the case of the lapwing goose, or the stickleback fish, it is clear that visual acuity is not an issue. Is this categorically different from humans experiencing "the scene in front of them?" It is hard to refrain from taking an anthropocentric perspective here. Most people, including scientists, apparently do take the anthropocentric perspective. Does this make principled sense? In my view, it categorically does not.

The animal interfaces are apparently successful, as evident from the fact that the species are "still here," even after evolutionary time spans. Evolution optimizes *utility* (often denoted "fitness"), not *veridicality*. Thus it may be said that *evolution breeds illusion*.[7] From the Marrean, or Bayesian point of view, this is not palatable.

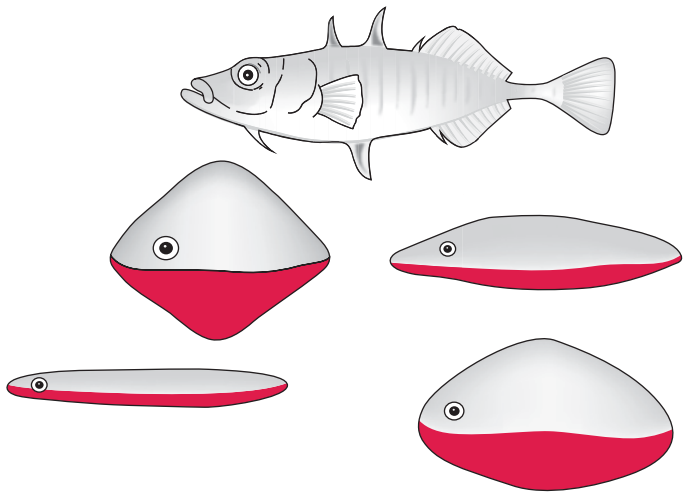


Figure I.8-5. Nico Tinbergen (1952). Fake stickleback male "models," preferred by the female over the real "he-man" fish. (An instance of "super-pornography" if there ever was one.)

Inverse optics will hardly turn a brick into an egg, given plenty of resolution and good views. The Bayesian prior for something near to the nest to be an egg must be close to 1, whereas the likelihood for something big, reddish, and squarish to be an egg must be close to zero. Thus the Bayesian principle is hardly useful here.

A major reason why these approaches are powerless is that they rely on the God's eye view. The ethologist, "user interface" concept is much more apt. But then the conventional notion of "illusion" can no longer be upheld.

One would think that the application of these ideas to the human condition would be the obvious way to go for a scientist. The mainstream appears to prefer a marked anthropocentric perspective (Grey, 1993), however. This strongly influences the way people conceive of "illusion."

The two limiting views lead to agreement on some, and opposite views on many, issues that are important in any discussion of visual illusions. A short list of such topics, by no means complete, certainly includes

- *Veridicality*. In the inverse optics, objectivist[8] view, perceptions are either veridical or erroneous. Veridicality simply means correct representation of physical reality. The symbols of *mentalese* (Fodor, 1975) derive their meaning from a preestablished correspondence with real objects. Objects are traced with ease because they are defined by unique subsets of all preestablished properties. It implies a God's eye view. Bizarre as this appears at first blush, this is actually a mainstream conviction. In contradistinction, in controlled hallucination veridicality is no more than an empty word. It is replaced with the notion of utility. Evolution optimizes utility, rather than veridicality. This implies that *qualia* (Schrödinger, 1958) are mental paint (Block, 2003), rather than

representations of physical properties, which is an incoherent notion anyway. These concepts are crucial in the discussion of illusions.

- *Embodiment*. There is general agreement that minds are embodied and that this embodiment is crucial to mental development. Meanings ultimately derive from uncontradicted experience in bodily sensorimotor actions. There is some agreement on the fact that sensorimotor actions do not necessarily imply consciousness. *Sensing* and *seeing* are categorically different. There is disagreement on whether all seeing is sensing, that is, whether seeing (in the sense of awareness) can properly be held to exist.[9] This is important in the discussion of illusions, especially if one confuses awareness with sensorimotor patterns, as is commonly the case in studies that "objectify" illusions.[10]
- *Neural substrate*. There is a widespread chaos of various disagreements on the meaning of neural activity. Inverse optics attaches meaning to any neural signal. For instance, some cells in the primary visual cortex are denoted "edge detectors." This presupposes the existence of edges before they are detected and labels their output as meaning. In contradistinction, for the controlled hallucination model there cannot be any "edge detectors" but only cells with certain receptive fields. Edges are hallucinated, and a probing of neural activity reveals a degree of belief that might be associated with such hallucinations. This is of frequent importance in the discussion of illusions, for example, in explanations involving lateral inhibition.
- *Local sign*. It is often taken for granted that visual awareness must have the same spatial structure as that of certain electrochemical processes in the brain. This echoes the early isomorphism ideas of the Gestalt school (Köhler, 1969), ideas that are generally ridiculed nowadays. Yet many mainstream researchers take it for granted (to offer just one example out of many possibilities) that it is of importance to awareness that the retina is a spherical surface with the anterior nodal point of the optics somewhere in between the anterior pole and the center. Such (often rather naïve) isomorphism notions play a role in many discussions of illusions. This ignores the crucial topic of local sign (Lotze, 1852) and implies some sort of homunculus mechanism.
- *Physiological bottlenecks*. There is general agreement that what fails to make the retinal receptor activity pattern will not possess the "causal power" to influence immediate visual awareness. Extremely fine detail will not make the retinal irradiance pattern, thus Jupiter's moons cannot be seen with the unarmed eye (Dutton, 1976). Likewise, people without eyes will be blind. Although few will dispute such examples, there is less and less agreement as one moves up into the visual system. For instance, the spatiotemporal pattern of photon absorptions is still present at the level of the cone synapses but probably lost at the level of the cortex, and so forth. Can it influence awareness? What does not make the visual stream cannot influence inverse optics results, nor be effective in pruning hallucinations. In the latter case, hallucination may "fill in" gaps; in the former, one is left with a problem.
- *Gestalt organization*. In the inverse optics scheme there is no room for such a thing as "Gestalt organization." In this view the Gestalt phenomena are interpreted as effects of singular, ambiguous, or insufficient data. The resulting representations are illusions because in error. The controlled hallucination scheme handles the Gestalt formation quite naturally, with the Gestalts being elements of the user interface. The resulting presentations reveal a reality that may differ from the realities obtained by touch or by vision after changing the vantage point. (See Fig. I.8-6.)

This short introduction must suffice as a general background. The bottom line is that our understanding of the overall structure of the microgenetic process is scant. We understand some of the substructures, mainly because of the bottlenecks they introduce, although no consensus has ever been reached as to their ontological status, nor of their precise place in the overall picture.

TAXONOMY

A short investigation on the Internet (simply use "visual illusion" in Google) yields numerous collections of diverse "visual" or "optical" illusions. There are dozens of instances to be found that are categorically distinct. As one soon discovers, this collection is really an odd lot! If there is any attempt at ordering at all, one finds such headings as "geometrical," "brightness," "color," "motion," and so forth. In some cases an attempt is made at a functional order, for instance, "amodal completion," "contrast," "assimilation," "after effect," and so forth. Some functions are indeed purely phenomenological, whereas others are principles borrowed from neurophysiology in thin disguise.

Given the sheer extent of this chaotic zoo, a more systematic taxonomy would be desirable. One with some kind of sequential, partial order, perhaps like a hierarchy, that might serve like a sort of "psycho-anatomy." However, such an ideal seems a long way off.[11]

The following list of a dozen or so categories is loosely structured with respect to the ontological levels, whereas they are usually "explained." I distinguish the ontologically distinct levels of physics, physiology, information theory, Gestalt formation, and microgenesis of visual awareness.

Physics

Optical Factors

As the ancients noticed (Lucretius c. 99 bc–c. 55 bc), objects at a distance are seen as round even when they

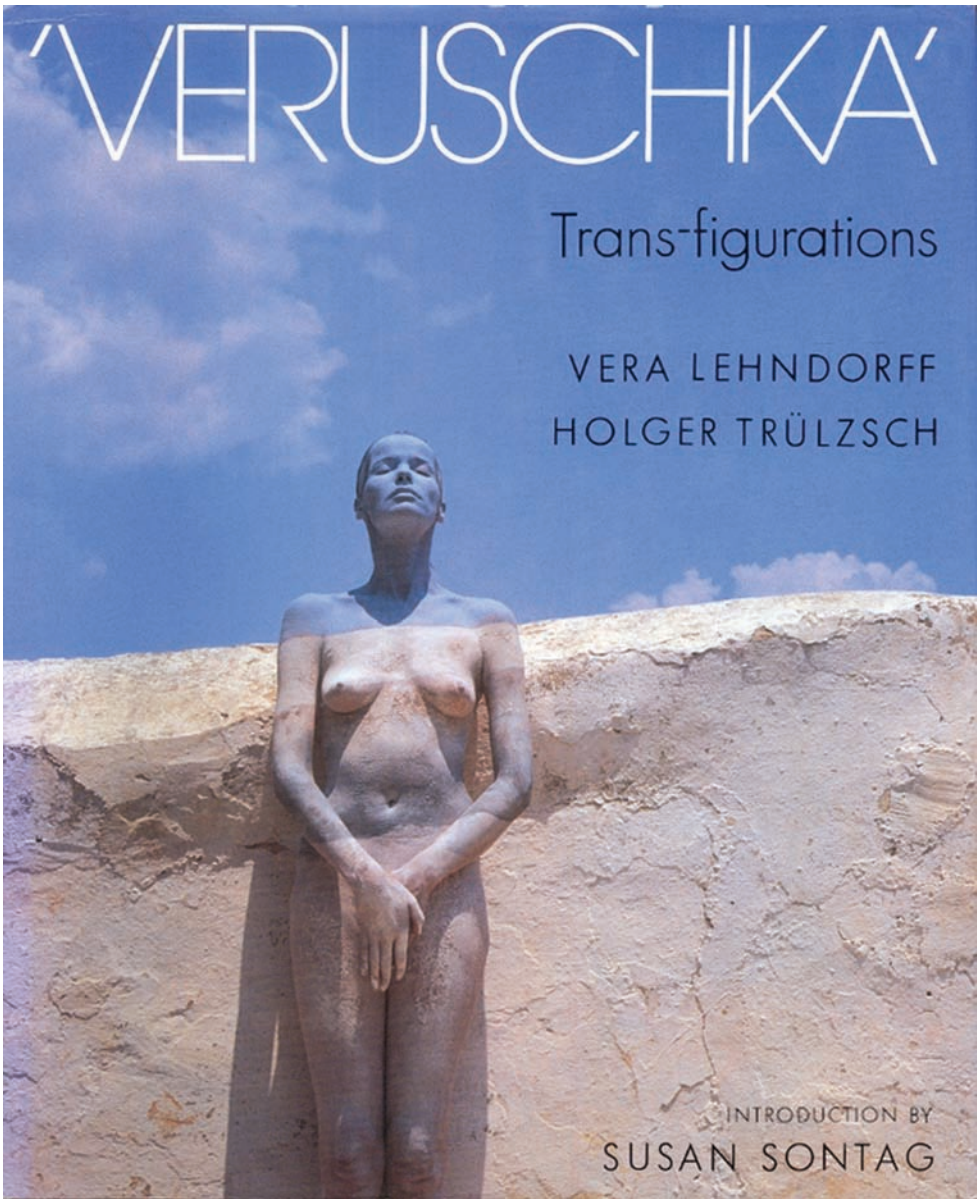


Figure I.8-6. *Veruschka: Transfigurations* (Lehndorff, Trulzsch, & Sontag, 1989; Vera Gräfin von Lehndorff-Steinort, b. 1939). What a little application of body paint can do! Of course the camera standpoint is crucial here. This is a very singular view. Small deviations from the viewpoint cause enormous changes in the presentation.



Figure I.8-7. *The Elements of Drawing* (Ruskin, 1857). A leaf seen at great distance appears to hover in the air. Is this an illusion? Certainly the stem is present in the environment, as leaves do not levitate. Thus the presentation fails to be "veridical."

are angular. The reason is the limited resolution of the eye's optics (Fig. I.8-7; Ruskin [1857] speaks of "mystery" when discussing the scale issue). Is this an illusion? It all depends on one's views. The physicist, unlike the philosopher, understands that *nothing* can be measured with infinite precision.[12] This is perhaps surprising in view of the fact that so many contemporary philosophers try to pose as physicists. Limited resolution is a basic fact of life (Florack, 1997). Arguments based on limited optical resolution have been used to explain simple geometrical illusions. One produces a low-resolution instance of the stimulus and shows that a certain summary description of it (such as a ridge of irradiance [Verstraelen, 2005]) deviates from the high-resolution stimulus in the same sense as the presentation.

Coarse Graining

Coarse graining has many effects that are often denoted as "illusion," although perhaps unfortunately so. Examples (Fig. I.8-8) are partitive color mixture (book printing



Figure I.8-8. Roy Lichtenstein (1923–1997), a painting composed of painted "Benday dots." The dots visually "mix" to uniform surfaces, even though the dots remain visible. The pleasure of viewing such a painting depends upon the presence of the dots being visually obvious. The painting needs to be "coarse" in order to work. (http://www.aaronartprints.org/lichtenstein-girlwithhairribbon.php)

depends on that) and apparent motion (TV and the movie industry exploit that). This notion of "illusion" is similar to Eddington's (1928) concept of a solid table that "in reality" consists largely of empty space.

Visual Space Effects

The stationary, monocular observer may group visual objects at very different depths as if they are close in the visual field (Hildebrand, 1893). Well-known examples are Alice's holiday snapshots of the tower at Pisa being pushed over by Bill (or vice versa; Fig. I.8-9). In cases of grouping of blobs in the visual field corresponding to objects that are far apart in the scene, one experiences the effects typical of camouflage.

Physiology

Lateral Inhibition

Haldan Keffer Hartline (1903–1983) received the Nobel Prize (1967) for his work on the lateral inhibition in *Limulus Polyphemus*. This triggered many explanations of visual illusions, the Mach band phenomenon to start with (Ratliff, 1965). Of course these "explanations" depend on a variety of isomorphisms, with all the problems this implies. (See Fig. I.8-10.)

Aftereffects

Aftereffects (Fig. I.8-11) occur on all levels and scales, in all submodalities. It is conventionally explained through adaptation of neural subsystems, although such explanations are by no means without their problems.

Information Theory: Structural Complexity

High structural complexity often suffices to hide things. Not being able to see things that are really there in plain sight is often denoted as illusion. It evidently is not; it merely reveals a bottleneck in the microgenetic process (Figs. I.8-12 and I.8-13). Structural complexity is important because the visual system is sensitive to even slight information overload. Complicated distributions appear as uniform textures, although the substructure changes from place to place. In depictions of a group of people the number of legs may not fit the number of persons, a fact that may be found by *counting* but cannot be *seen*. Spatiotemporal textures appear as scintillations of various kinds.

Gestalt Formation

Grouping

Various grouping phenomena in the visual field lead to illusions that are often referred to as "camouflage." Camouflage is often indicated as a kind of illusion, whereas it is more typically a form of deception. Grouping of random configurations may lead to a variety of diverse presentations (Fig. I.8-14). Grouping in space-time yields illusions of motion or continuous deformation.

Ambiguity

All pictures are ambiguous (Fig. I.8-15). Some pictures are intentionally designed to allow just two instant presentations that are mutually exclusive. One of the simplest examples is the Necker cube; more complicated is Jastrow's (1899) famous duck-rabbit. Are these to be considered illusions? Probably not, although in the former the presentation of a cube (any of the two common interpretations) instead of a planar figure might be construed as such. For instance, the three-dimensional cube in the visual presentation evidently conflicts with haptic exploration.

Amodal Completion

Amodal completion, or "filling in the holes," occurs on many different levels and scales (Fig. I.8-16). Although there are



Figure I.8-9. Man kicking over the leaning tower of Pisa. The boot and the tower are at different depths, but depths are not represented by location in the picture plane. The meeting of the boot and the tower looks hardly "accidental." Microgeny groups tower and boot in the presentation of a "coherent" event. Such a process is likely to have great utility in daily life; one should be grateful for it. (http://coda-craven.org/man-pushing-the-tower-4k-full-uhd-desktop-wallpaper)



Figure I.8-10. Georges Seurat (1859–1891), *Study for Bathers at Asnieres*. Seurat puts lateral effects to good use. He has only black on white paper. Notice the full roundness of the torso, due to—among more—the treatment of the background. The range of tones available in this technique is very low; Seurat makes surprisingly good use of it, artfully composing a drawing (figure plus background) to boot.



Figure I.8-11. Giacomo Balla (1871–1958), *Dog on a Leash*. Notice the effect of after images. No doubt these futurist images were inspired by the experience with early movies (Marey, 1890).



Figure I.8-12. Hans Makart (1840–1884), *Sketch of the Arrival of Charles V at Antwerpen*. This sketch evidently overfills the visual buffer. How many people are there? The final painting is huge, and one needs considerable time to view it thoroughly. This sketch is perhaps even more challenging. Notice how our presentations appear to contain groups of people that are not individualized but somehow merge into a "crowd-texture."

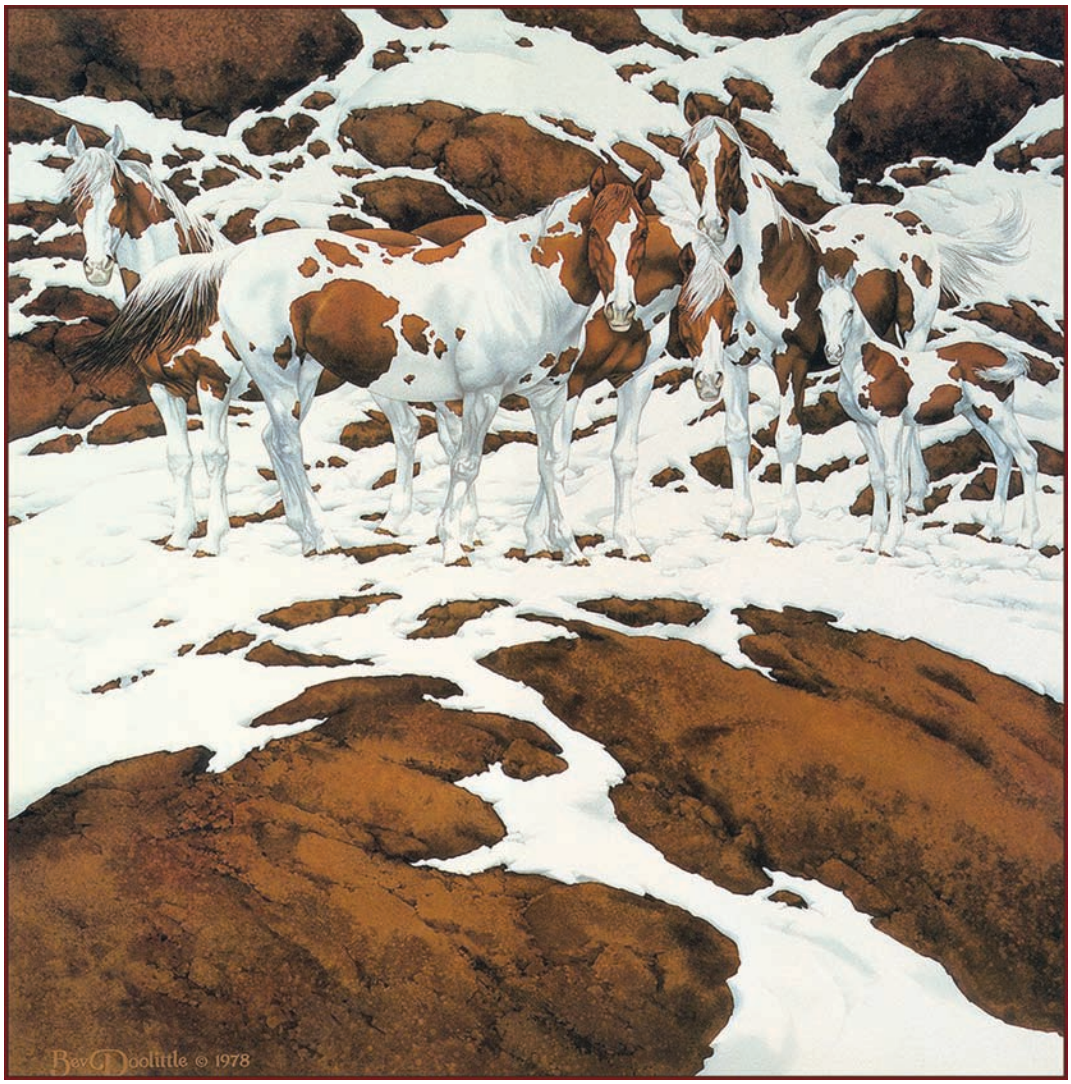


Figure I.8-13. Bev Doolittle (b. 1947), *Pintos*. How many horses are there? Yet they are all in plain sight, painted in a realistic fashion. Although the sheer number of horses might fit the mind's bottlenecks, the fact that one must patch them together from patches that have to be identified actually exceeds the power of most people's microgenetic processes. (http://www.greenwichworkshop.com/details/default.asp?p=19&t=1)

many attempts at bottom-up explanations, such phenomena fit much more naturally in the controlled hallucination view. A key example is the Kanizsa (1955) triangle, but the apparent "regularization" of disorderly patterns may also be understood as an amodal completion, this time the positing of a regular frame. Similarly, the Craik-O'Brian-Cornsweet illusion (Cornsweet, 1970) can be understood as an amodal completion, since the absolute brightness levels do not survive the visual front end. "Completion" may be forced, as when the relevant structure does not make it into the visual stream or has to be freely invented, as in the case of the Kanizsa triangle.

Visual Field Effects

Many effects involve lateral interactions in simple layered networks. Examples are contrast, assimilation, filling-in, and so forth (Fig. I.8-17). In this case "contrast" is similar to "lateral inhibition" but involves no neural substrate. It is often suggested that simple "neural networks" explain such effects. It is usually unclear where such networks are supposed to be situated, and it is probably best to understand the layered networks as formal devices. When understood as actual neural networks this once again depends on a variety of isomorphisms. More important, modern instances of contrast and assimilation show that certain "higher order" interpretations are often involved (Adelson, 2000), simple layered networks being insufficient.

Visual Space Effects

In many cases flat patterns enter awareness as three-dimensional configurations. It may actually be hard to experience the patterns as flat (Fig. I.8-18). This seems to be due to microgenetic factors close to Gestalt formation, and they form an important fundament of daily vision. Whether one interprets this as "illusion" is a matter of

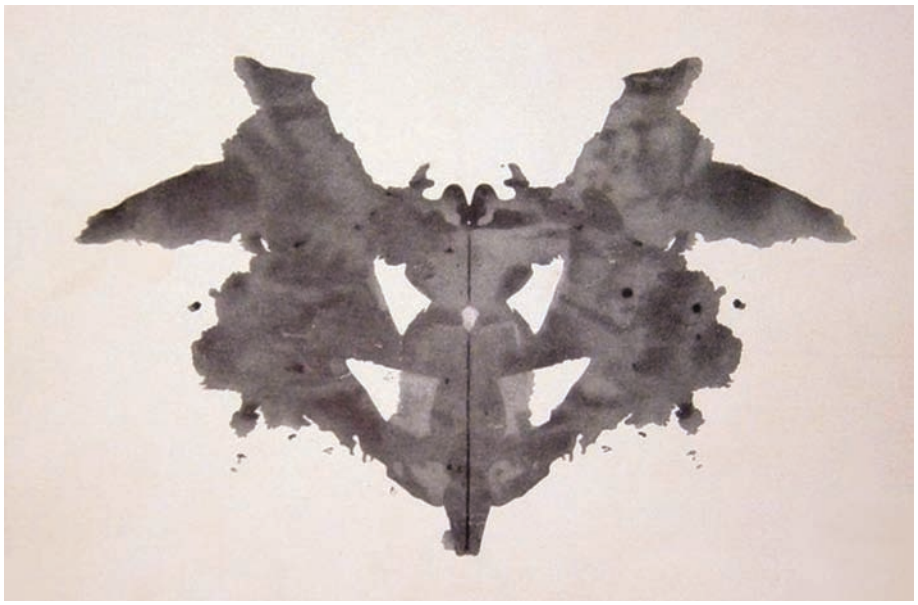


Figure I.8-14. An item from the Rorschach test (1927). Suppose Alice sees a "witch with a machine gun" and Bill a "baby angel with clipped wings." Who is to call them wrong? If one protests that both "should see" an ink blob, then one should apply the same test to Leonardo's *Mona Lisa*: Is that just pigments applied to a planar substrate? What if Chris sees a "flying pig" in it? Remember that Freud (1910) saw a vulture in Leonardo's *The Virgin and Child with St. Anne*. Was he mad? If so, then all of us are, at least occasionally.



Figure I.8-15. Left: Giuseppe Arcimboldo (1527–1593), *Librarian*. A planar surface covered with pigments, an odd pile of books, a stuffy old librarian, or perhaps something else? (http://www.the-athenaeum.org/art/list.php?m=o&s=du&oid=502.&f=a&fa=4085). Center: The billowing folds of Gianlorenzo Bernini's (1598–1680) *Neptune* show the head of a dolphin if viewed from a certain angle. There is little doubt that this dolphin was intentional. Is it an illusion? If so, of what type? (https://nonvisedvoce.wordpress.com/category/sculpture/). Right: Salvador Dali (1905–1989) with a composition based on a number of nude female models. Most people see a skull in this strange heap of bodies. (https://archishadow.wordpress.com/philippe-halsman-2/)

taste. We certainly would not like to go without such illusions, as they have considerable utility.

Microgenesis of Visual Awareness

Construction of Space-Time

As various glimpses are integrated in a glance, the glimpses are typically less than perfectly registered, leading to a variety of scintillating and apparent movement effects. There does not seem to be a strict global topology; one has a locally disorderly awareness, often changing from one presentation to the next (Fig. I.8-19).

Pictorial Space

Pictures are planar distributions of pigments or radiance. Yet they are often confused with physical objects or scenes. For example, we see a portrait (a face) instead of a complicated gray-tone distribution, or a certain arrangement of



Figure I.8-16. Felix Valloton (1865–1925), *The Money*. Valloton uses amodal completion. Notice the right-side contour of the man explaining the monetary problem to the woman. Is it there? In our presentations the man is certainly not "incomplete." But is there the illusion of a right-side contour? What would make this case different from the familiar Kanizsa triangle? It does not have contours either.

black lines on white paper (e.g., a "smiley face" or "**(-:**"). If someone subscribes to the inverse optics model—that pictorial space is an illusion—he or she should try to grab a pictorial object! If one subscribes to the controlled hallucination model, it is perhaps not all that remarkable (Fig. I.8-20).

The Scene

This will be the primordial illusion (Fig. I.8-21) for those who subscribe to the controlled hallucination model. If one is of the inverse optics creed, things depend. If the observer is a brain in a vat, being provided with fake "optical data," all is illusion. If the observer is in the presence of others, who report seeing similar things, the observer may be inclined to rate his or her seeing as veridical. Phenomenologically there would be no difference. *Veridicality exists only in the perception of others.* This is a key observation. In the first-person account, veridicality remains undefined.[13]

Some items, such as coarse graining, may be placed into different categories. This is a sign that they are more like general principles that cut through the ontological levels. Such principles are *structural complexity, information capacity, coarse graining*, and *ambiguity*. The *Gestalt principles* may also be applied on various levels.

The list seems to be more or less complete and to be made up of more or less independent and more or less nonredundant categories. This is perhaps a start, as any classification should at least satisfy the properties of completeness, mutual independence, and nonredundancy. We regret that the list is apparently not easily organized on some principled basis though.

An almost flat list of classes is as far from a proper taxonomy as it can get. Several groupings can be made on the basis of different criteria. The one I applied here is based on the ontological level of explanation. Perhaps this is all that can be done; perhaps the zoo of illusions *is* an odd lot. Of course it remains possible that different categories, cutting right through the ones proposed here, may turn out to be more useful.

PSYCHOANATOMY

The use of illusions as tools in vision research was pushed forcefully by the Gestalt school (Metzger, 1936). Many of their strongest arguments are *ad oculos* demonstrations that are absolutely compelling and can hardly be argued with. The hope is that illusions can be used as tools (hopefully scalpels and needles instead of sledge hammers and shovels) in a psychoanatomy of vision.



Figure I.8-17. On this historical (Willem Blaeu [1571–1638]) map of Europe, the nations are distinguished from each other by color. The color is applied near the boundaries but visually "spreads" over the regions. (Atlas Blaeu-Van der Hem Nationalbibliothek, Vienna, Austria.)

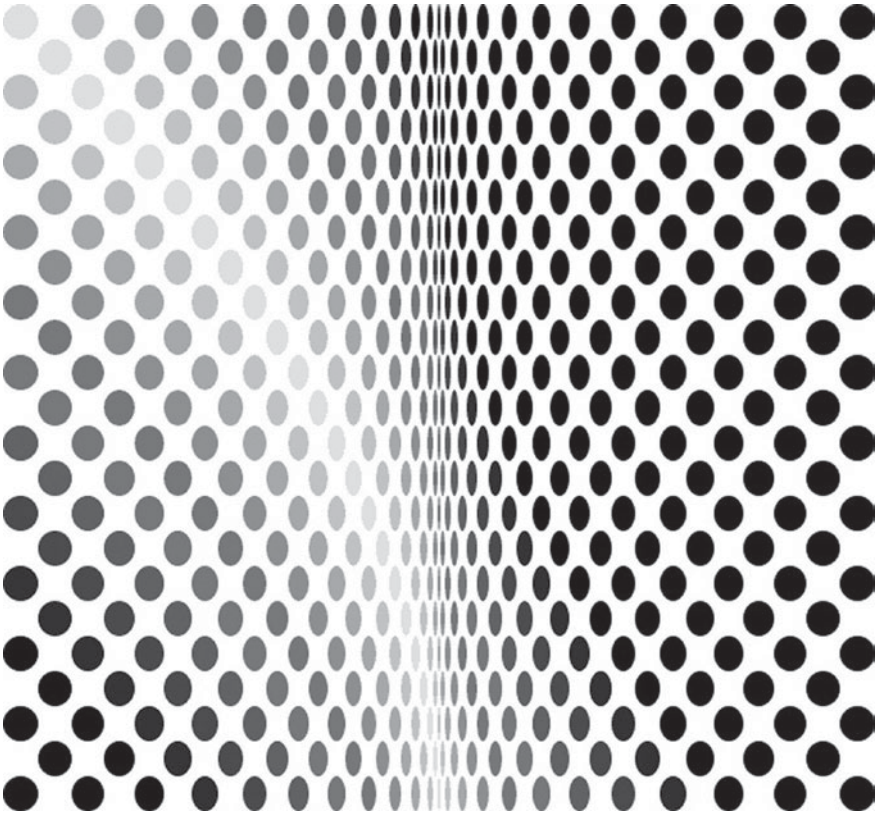


Figure I.8-18. Bridget Riley's (b. 1931) op-art shows an array of elliptical polka dots of various gray tones. It is hard not to see three-dimensional modulations of the picture plane.

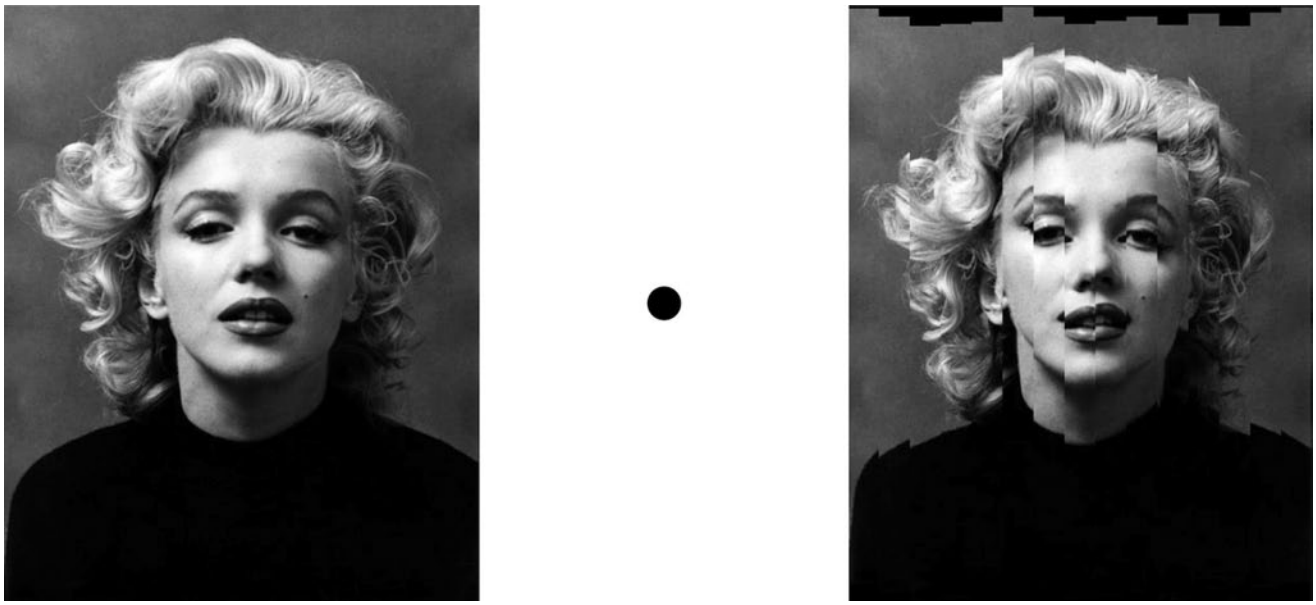


Figure I.8-19. The figure at the right was obtained by cutting the original in narrow vertical strips, shifting the strips vertically by rather large amounts, and reassembling. In focal vision one sees some "noise"; in preattentive vision the portrait looks normal. The construction of space is not particularly bothered by the local disarray. (Fixate on the black center dot to compare the two images.) (Photo by ben ross, 1953.)



Figure I.8-20. Umberto Boccioni (1882–1916), *The City Enters the Window*. Obviously a *picture*, yet a fully convincing three-dimensional space. It is no doubt a different space from one we experience in daily life, so the "veridicality" of the painting qua representation is not great. This qualifies it as an "illusion" in many books. Located at: Sprengel Museum, Hanover, Germany.

The disciples of the inverse optics model would not hesitate to map the psychoanatomy on the neuroanatomy (Poggio, 1990). For the adherents of the controlled hallucination model, such a move is rather less immediate. Because we really do not know how neural activity might "translate" into awareness (Leibniz's [1714/1840] psychophysical parallelism), the only trustworthy principle for the sciences is that

> *what is not in neural activity cannot influence the structure of awareness.*

That is to say, one exploits partial blindnesses (Lissauer, 1890), or *agnosias*. This has yielded many interesting results, exploiting either brain lesions or sophisticated psychophysical techniques.

Given the fact that one's identification of phenomenological bottlenecks with neurophysiological ones depends crucially on one's initial convictions, it seems prudent to postpone such identifications of psychoanatomy with neuroanatomy, at least in the majority of cases. Such practices have certainly not yielded a generally accepted overall, detailed picture thus far. It seems unlikely that this will significantly change in the near future.

An apt use of illusions may well serve as a kind of "instant psychophysics" (Richards, Koenderink, & Hoffman, 1985) approach to selective blindnesses in immediate visual awareness. In my view, this is where the work on "illusions" really shines and where the most promising approaches for future developments are found. This is unashamedly *psychoanatomy*, rather than *neuroanatomy,* although one hardly dares to openly come right out and say this these days.[14] The structure of the microgeny of immediate visual awareness is still very far from being understood. But even if that were in place, the translation of an understanding of microgenesis into neurocognitive terms poses another, almost unsurmountable hurdle, the possibility (or even the very meaning) of which will not even be granted by some. In my view a certain amount of reserve, or modesty, is appropriate here.

AUTHOR'S NOTE

Unfortunately many readers will be handicapped by limited linguistic resources provided in the References, as an appreciable fraction of original literature is in German, French, Italian, Russian, and so on, rather than English. I have nevertheless cited primarily original sources. The way to proceed is to use the Internet, starting from the cited source. In virtually all cases this will yield references to various translations, often complete texts, or a precis in English.

ACKNOWLEDGMENTS

I am grateful to Andrea van Doorn and Johan Wagemans for support and critique. This work was supported by the Methusalem program by the Flemish Government (METH/08/02), awarded to Johan Wagemans.

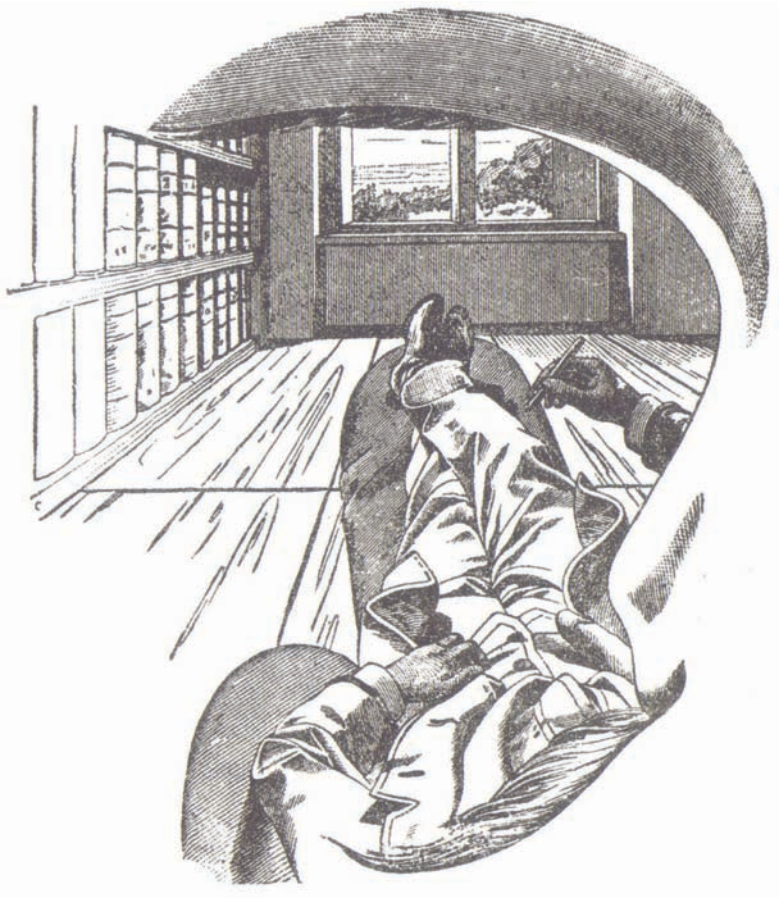


Figure I.8-21. Ernst Mach (1886). The illusion of the scene. According to Mach: "Thing, body, matter, are nothing apart from the combinations of the elements,—the colours, sounds, and so forth—nothing apart from their so-called attributes."

NOTES

1. Jack London (1989), in the autobiography *John Barleycorn,* describes such a person as "the man whom we all know, stupid, unimaginative, whose brain is bitten numbly by numb maggots; who walks generously with wide-spread, tentative legs, falls frequently in the gutter, and who sees, in the extremity of his ecstasy, blue mice and pink elephants. He is the type that gives rise to the jokes in the funny papers."
2. As a simple market girl Saint Catharine of Sienna (b. 1347) saw Christ on a throne in the sky. Whereas she was probably the only one to make this remarkable observation in the middle of a crowd, it was (historically, and perhaps remarkably) not recorded as a "hallucination" but as a "vision." We do not further explore the topic of visions alone in this chapter.
3. Such an entity is precategoric, so "rope-snake-like" merely indicates a certain coherent substructure of visual awareness; it does in no way refer to any animals or utensils.
4. A "format" indeed bestows meaning on mere structure. The same structure (a few key strokes to a laptop, say) means different things to different programs (a password, a number, a string, a command, nonsense, etc.) because of this.
5. In the game of 20 Questions, one person decides on a certain word, and another person has to guess what the first person has in mind. The second person is allowed 20 questions of a binary nature ("this" or "that"). The second person usually "wins." Most of the fun is in watching the tortuous route (starting as a random walk) toward the solution. In the gist game, an observer plays the game against nature.
6. Notice that this is a very interesting issue. "Illusion" is conventionally understood in a categorial sense. *Something is an illusion, or it is not.* Thus to allow "degrees of illusion" implies a complete rethinking.
7. This is another potentially very important issue to ponder. The mainstream notion is—of course—that evolution breeds veridicality.
8. Here I mean "metaphysical objectivism," rather than the philosophy of Ayn Rand (1943/1966).
9. This should surprise readers who are (perhaps unknowingly) convinced of the fact that they have visual awareness (zombies happily avoid this problem by not experiencing at all). If they are stubborn enough to stick to the notion that they are aware, they should perhaps reconsider their status as scientists.
10. Reconsider the documented phenomenon of "blind sight" in this connection.
11. The late Richard Gregory (d. 2009), surely a top expert in these matters, published a book that attempts an ordering of illusions. However, I am unable to interpret it as a logical structure.
12. A common example of this failing to understand the basics of measurement is human spectral sensitivity. The retinal machinery samples the radiant power in three bands, thus measuring a low-resolution spectrum as compared to some physical devices. This is often construed as that we "do not see by wavelength." Neither does any physical apparatus, as they all have finite resolution. It makes no essential difference whether the number of degrees of freedom is 3, 10, 100, 1,000 (the ball park of quality equipment), or more. Even dichromatic retinæ sample spectra. Limited resolution is no reason to hold that "color is mental paint"; the same would hold for the best spectrographic equipment.
13. This is why people rarely notice they are hallucinating. The only thing people may notice is the multimodal coherence of their awareness and the degree to which their awareness agrees with reports by others. That "veridicality" exists

only in the perception of others is crucial, though rarely acknowledged.

14. Research grants are certain to go to the neuro-something applications nowadays, regardless of whether the claims make sense or not. Anything psycho-something is likely to end up on the bottom of the pile: "grant if resources allow." Perhaps unfortunately, this is a mondial trend, unlikely to pass over soon. After all, this involves the taxpayer's money.

REFERENCES

Adelson, E. H. (2000). Lightness perception and lightness illusions. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (2d ed., pp. 339–351). Cambridge, MA: MIT Press.

Berkeley, G. (1709). *An essay towards a new theory of vision,* Dublin: Aaron Rhames.

Block, N. (2003). Mental paint. In M. Hahn & B. Ramberg (Eds.), *Reflections and replies: Essays on the philosophy of Tyler Burge* (pp. 165–200). Cambridge, MA: MIT Press.

Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence, 47*(1–3), 139–159.

Brown, J. W. (1977). *Mind, brain and consciousness.* New York: Academic Press.

Clark, A. (2004). Feature-placing and proto-objects. *Philosophical Psychology, 17*(4), 451–477.

Cornsweet, T. N. (1970). *Visual perception.* New York: Academic Press.

Deussen, P. (1883). *Das System des Vedânta nach den Brahma-Sûtra's des Bâdarâyana und dem Commentare des Çankara über dieselben als ein Compendium der dogmatik der Brahmanismus vom Standpunkte des Çankara aus.* Leipzig: F. A. Brockhaus.

Dutton, D. (1976). Naked-eye observations of Jupiter's moons. *Sky and Telescope* (December): 482–484.

Eddington, A. S. (1928). *The nature of the physical world.* New York: MacMillan.

Florack, L. (1997). *Image structure.* Dordrecht: Kluwer Academic.

Fodor, J. A. (1975). *The language of thought.* New York: Crowell Press.

Forsyth, D. A., & Ponce, J. (2003). *Computer vision: A modern approach.* Upper Saddle River, NJ: Prentice Hall.

Freud, S. (1910). *Eine Kindheitserinnerung des Leonardo da Vinci.* Vol. 8. Leipzig: F. Deuticke.

Grey, W. (1993). Anthropocentrism and deep ecology. *Australiasian Journal of Philosophy, 71*(4), 463–475.

Gregory, R. L. (2009). *Seeing through illusions,* Oxford: Oxford University Press.

Gregory, R. L., & Heard, P. (1979). Border locking and the café wall illusion. *Perception, 8*, 365–380.

Hevelius, J. (1690). *Firmamentum Sobiescianum sive Uranographia.* Gdansk, Germany: Johannis Zachariae Stollii.

Hildebrand, A. (1893). *Das problem der form in den bildenden kunst.* Strassbourg: Heitz & Mündel.

Hoffman, D. (2009). The interface theory of perception: Natural selection drives true perception to swift extinction. In S. Dickinson, M. Tarr, A. Leonardis, & B. Schiele (Eds.), *Object categorization: Computer and human vision perspectives* (pp. 148–165). Cambridge, UK: Cambridge University Press.

Hume, D. (1975). *Enquiries concerning human understanding and concerning the principles of morals.* Oxford: Oxford University Press. (Original work published 1777)

Jastrov, J. (1899). The mind's eye. *Popular Science Monthly, 54*, 299–312.

Kanizsa, G. (1955). Margini quasi-percettivi in campi con stimolazione omogenea. *Rivista di Psicologia, 49*(1), 7–30.

Knill, D., & Richards, W. (Eds.). (1996). *Perception as Bayesian inference.* Cambridge UK: Cambridge University Press.

Koenderink, J. J. (2011). Information. In L. Albertazzi, G. van Tonder, & D. Vishwanath (Eds.), *Perception beyond inference: The information content of visual processes* (pp. 345–390). Cambridge, MA: MIT Press.

Köhler W. (1969). *The task of Gestalt psychology.* Princeton, NJ: Princeton University Press.

Kuhn, G., Amiani, A. A., & Rensink, R. A. (2008). Towards a science of magic. *Trends in Cognitive Science*, *12*(9), 349–354.

Lakoff, G. (1987). *Women, fire, and dangerous things.* Chicago: University of Chicago Press.

Lehndorff, V., Trulzsch, H., & Sontag, S. (1989). *Veruschka: Transfigurations.* London: Thames & Hudson.

Leibniz, G. (original 1714, 1840). The Monadology, English translation by Robert Latta 1898, links, scalable text and printable version. Downloadable as pdf, doc or djvu files. http://home.datacomm.ch/kerguelen/monadology/

Lissauer, H. (1890). Ein Fall von Seelenblindheit nebst einem Beitrag zur Theorie derselben. *Archiv für Psychiatrie und Nervenkrankheiten*, *21*, 222–270.

Locke, J. (1690). *An essay concerning human understanding.* London: Eliz. Holt.

London, J. (1989). *John Barleycorn.* New York: Oxford University Press. (Original work published 1913)

Lorenz, K. (1973). *Die Rückseite des Spiegels: Versuch einer Naturgeschichte menschlichen Erkennens.* München: R. Piper.

Lotze, H. (1852). *Medicinische Psychologie oder Physiologie der Seele.* Leipzig, Germany: Weidmann'sche Buchhandlung.

Lucretius. On the Nature of Things Written 50 B.C.E. Translated by William Ellery Leonard. http://classics.mit.edu/Carus/nature_things.html

Mach, E. (1886). *Die Analyse der Empfindungen und das Verhältnis der Physischen zum Psychischen*, 2. verm. Aufl. Jena: Fischer, 1900. [Die erste Auflage erschien 1886 u.d.T.: Beiträge zur Analyse der Empfindung]

Marey, É.-J. (1890). *Le Vol des Oiseaux.* Paris: G. Masson.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information.* New York: Freeman.

Metzger, W. (1936). *Gesetze des Sehens Erstauflage* 2. erweiterte Auflage. Frankfurt: Verlag Waldemar Kramer.

Nico Tinbergen (1952). The Curious Behavior of the Stickleback Scientific American, 187(6), 22-26.

Oliva, A., & Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research: Visual Perception, 155*, 23–36.

Penrose, R. (1999). *The emperor's new mind.* Oxford: Oxford University Press. (Original work published 1989)

Poggio, T. (1990). A theory of how the brain might work. *Cold Spring Harbor Symposia on Quantitative Biology, 55*, 899–910.

Poli, R. (2006). Levels of reality and the psychological stratum. *Revue Internationale de Philosophie, 61*(2), 163–180.

Rand, A. (1996). *The fountainhead* (50th anniversary ed.). New York: Signet. (Original work published 1943)

Ratliff, F. (1965). *Mach bands: Quantitative studies on neural networks in the retina.* San Francisco, CA: Holden-Day.

Rensink, R. A. (2002). Change detection. *Annual Review of Psychology*, *53*, 245–277.

Rensink, R. A. (2004). Visual sensing without seeing. *Psychological Science, 15*, 27–32.

Richards, W. A., Koenderink, J. J., & Hoffman, D. D. (1985). MIT AI-lab A.I. Memo No. 840.

Riedl, R. (1980). *Biologie der Erkenntnis: Die stammesgeschichtlichen Grundlagen der Vernunft.* Berlin: Parey.

Rorschach, H. (1927). *Rorschach test: Psychodiagnostic plates.* Cambridge, MA: Hogrefe.

Runge, Ph. O. (1810). *Farben-Kugel oder Construction des Verhältnisses aller Mischungen der Farben zu einander, und ihrer vollständigen Affinität, mit angehängtem Versuch einer*

Ableitung der Harmonie in den Zusammenstellungen der Farben. Hamburg: Friedrich Perthes.

Ruskin, J. (1857). *The elements of drawing, in three letters to beginners*. London: Spottiswoode.

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences, 9*, 16–20.

Schrödinger, E. (1943, February). What is life? The physical aspect of the living cell. Paper presented at the Dublin Institute for Advanced Studies at Trinity College, Dublin. http://whatislife.stanford.edu/LoCo_files/What-is-Life.pdf

Schrödinger, E. (1958). The mystery of sensual qualities. In E. Schrödinger, *What is life? With mind and matter and autobiographical sketches* (Cantro ed.) (pp. 153–164). Cambridge, UK: Cambridge University Press.

Tinbergen, N. (1951). *The study of instinct*. Oxford: Clarendon Press.

Uexküll, J. von (1909). *Umwelt und Innenwelt der Tiere*. Berlin: J. Springer.

Verstraelen, L. (2005). A geometrical description of visual perception. *Kragujevac Journal of Mathematics*, *28*, 7–17.

Watson, J. B. (1930). *Behaviorism* (rev. ed.). Chicago: University of Chicago Press.

Yuille, A., & Kersten, D. (2006). Vision as Bayesian inference: Analysis by synthesis? *Trends in Cognitive Sciences, 10*(7), 301–307.

Zöllner, K. F. (1860). Über eine neue Art von Pseudoskopie und ihre Beziehungen zu den von Plateau und Oppel beschrieben Bewegungsphænomenen. *Annalen der Physik 186*, 500–525.

Chapter 9
Why the Concept of "Visual Illusions" Is Misleading

Dale Purves, William T. Wojtach, and R. Beau Lotto

INTRODUCTION

Light projected onto the retina is unable to specify the variables in the physical world that we and other animals must respond to behaviorally (Fig. I.9-1). As a result, generating useful perceptions and behavioral responses on the basis of vision is difficult to understand. An explanation supported by recent studies of the way humans see lightness, brightness, color, geometry, and motion is that humans and other visual animals contend with this challenge by incorporating the accumulated consequences of trial-and-error behavior into visual system circuitry over evolutionary and individual time on a wholly empirical basis (Howe & Purves, 2005; Purves & Lotto, 2011; Purves, Williams, Nundy, & Lotto, 2004; Purves, Wojtach, & Lotto, 2011; Purves, Morgenstern, & Wojtach, 2015).

Since the late 1950s, the focus of visual neuroscience has been on the receptive field properties of neurons in the primary and higher-order visual pathways in experimental animals (Hubel & Wiesel, 2005; Palmer, 1999). Implicit in this research is the idea that perception arises from neural mechanisms that detect the features of retinal stimuli, analyze them hierarchically and in parallel, and eventually recombine this information in the visual cortex to represent the external world. However, given the inherently uncertain sources of retinal images and the inability of biological vision to measure reality, this paradigm is not feasible; representing the world via the features of retinal images would not lead to successful behavior. Therefore, the visual system evidently relies on a fundamentally different strategy that requires a shift in our understanding of perception, the purposes being served by visual system circuitry, and, by the same token, a different interpretation of "visual illusions."

A WHOLLY EMPIRICAL CONCEPT OF VISION

Although it is often assumed that the purpose of vision is to correctly represent the world, there is overwhelming evidence that this is not the case (Howe & Purves, 2005; Purves & Lotto, 2011; Purves et al., 2004; Purves et al., 2011). Whatever else this evidence suggests, it demonstrates that perceptions do not have to align with measured reality to be useful. If the aim of vision were to see things as they really are, then seeing them differently would indeed be to see illusions. But as described in the following section, the discrepancies between objective measurements and subjective experience are simply the signature of a strategy that has been biologically successful. While no one would presumably argue with the idea that vision evolved to promote useful behavior, this perspective explains the qualities of perception by the history of reproductive success rather than by feature analysis, Bayesian inference, or complex neural computations. Because of the importance of past behaviors on perception, we refer to this strategy as "wholly empirical."

EXPLAINING VISION IN AN EMPIRICAL FRAMEWORK

The biological principles underlying this concept of vision are straightforward. Natural selection would have instantiated random changes in the structure and function of the visual systems of ancestral forms according to how well the ensuing behavior served reproductive success. Any configuration of neural circuitry that better served the empirical link between visual stimuli and behavior that was successful in these terms would tend to increase among the members of the species, whereas less useful circuit configurations would not. Similarly, experience-dependent refinements of connectivity during postnatal development and adult life would allow individuals to contend with the challenge presented by ambiguous information more successfully than could be achieved on the basis of inherited circuitry alone. By progressively modifying the characteristics of visual stimuli created by primordial eyes and the associated neural connectivity according to reproductive success, the inability of vision to measure the world could be circumvented in favor of useful perceptions and behaviors (Purves, Monson, Sundararajan, & Wojtach, 2014). These gradual changes of connectivity in phylogeny and further adjustments during ontogeny provide two different mechanisms (natural selection vs. neural plasticity) and time frames (eons vs. minutes, days, or months) for instantiating the consequences of experience in the circuitry of the visual system.

The merits of this empirical framework in explaining perception can be assessed using natural scene databases

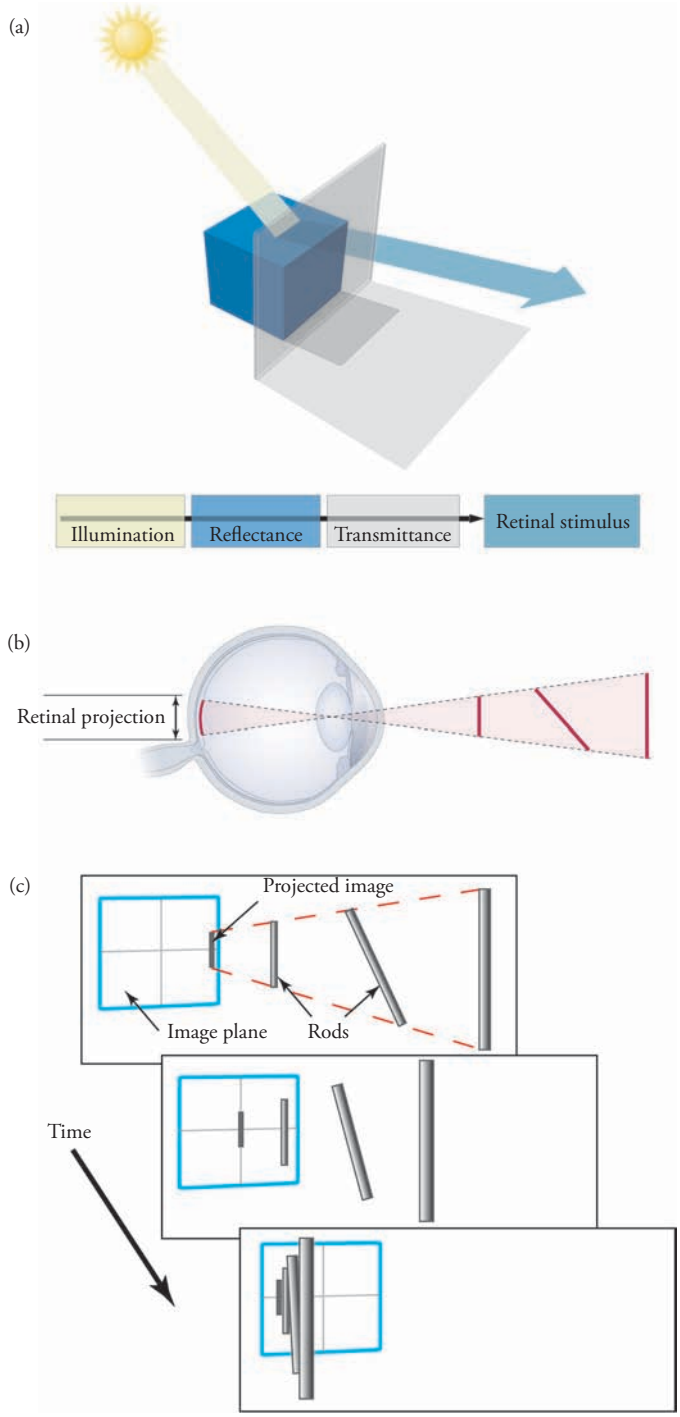


Figure I.9-1. The inherent ambiguity of the visual stimuli. (a) The conflation of illumination, reflectance, and transmittance in retinal images. Many combinations of these physical characteristics of the world can generate the same retinal stimulus. (b) The conflation of physical geometry in images. The same image can be generated by objects of different sizes, at different distances from the observer, and in different orientations. (c) The conflation of speed and direction in images of moving objects. The same projected motion on the retina can be generated by different objects, with various orientations, moving in different directions, and at different speeds in the physical world. (Based on Purves, Wojtach and Lotto, 2011.)

as proxies for the trial-and-error experience that, over time, would link biologically determined stimuli to useful visually guided behaviors, thereby contending with the inabilty of the visual system to measure the world (Purves & Lotto, 2011; Purves et al., 2014; Purves et al., 2011). The simplest example to demonstrate this is the way we respond to luminance. Since increasing the luminance of a stimulus increases the number of photons captured by photoreceptors and thus the relative action potential output of the retina at any given level of ambient light, a common-sense expectation is that physical measurements of light intensity and perceptions of lightness/brightness are directly proportional and that objects returning the same amount of light to the eye should appear equally light or bright. Perceptions of lightness and brightness, however, fail to meet these expectations. For instance, two patches having the same luminance are perceived as being differently light or bright when their backgrounds have different luminance values (Fig. I.9-2a). Moreover, presenting luminance targets in more complex contexts makes these effects much stronger (Figure I.9-2b; Gelb, 1929/1938; Gilchrist, 2009; Purves & Lotto, 2011).

To test the validity of this concept of vision, templates similar to the center-surround patterns in Figure I.9-2a can be used to sample a database of images arising from natural scenes (Fig. I.9-3). By tallying the frequency of occurrence of target luminance values in different surrounds, these data can indicate the historically successful behavior that would have linked the creation of biologically determined stimuli and reproductive success. Thus the identical targets in Figure I.9-2 appear differently light because the frequency of occurrence of the stimulus patterns will not have been the same in the two cases. By assigning lightness values in this way, perception will have tracked reproductive success, not the physical properties of the world.

For example, consider the targets (T) in Figure I.9-3a in terms of all the possible target luminance values in images from natural scenes that are surrounded by contexts with different luminance values. The probability distribution of target luminance values co-occurring with surrounding contexts having two different luminance values is shown in Figure I.9-3b. To predict the perceived lightness elicited by target luminance values on empirical grounds, the distribution of co-occurring target and surrounding luminance values can be expressed in cumulative form to define the summed probability that a particular target luminance will have occurred in a given surround over human experience (Fig. I.9-3c; note that ranking stimuli in this way avoids the fact that the same probability in the probability density function in Fig. I.9-3b specifies two luminance values). Thus the perception of lightness is predicted by the relative percentile rank (probability • 100%) of each co-occurring target and surround in these cumulative distributions, a higher rank corresponding to a greater perceived lightness of the target relative to the surround. Again, the rationale for predictions made on this basis is the inclusion of all factors relevant to successful behavior

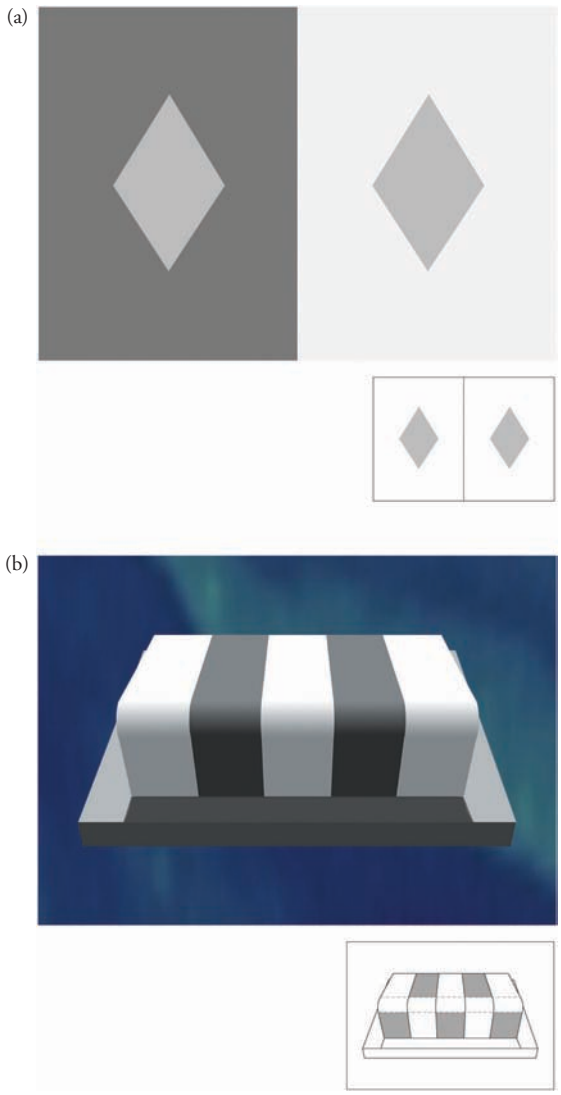


Figure I.9-2. Discrepancies between luminance and perceptions of lightness. (a) Standard demonstration of the simultaneous lightness contrast effect. A target (the diamond) on a less luminant background (left) is perceived as being lighter than a target on a more luminant background (right), even though the two targets have the same measured luminance; if the same targets are presented on a background with uniform luminance, they appear to have the same lightness (inset). (b) Although the target patches again have the same luminance and appear the same in a neutral setting (inset), a striking perceptual difference in lightness is generated by contextual information that one set of patches is in shadow (those on the riser of the step) and another set in light (those on the surface of the step). (Based on Purves & Lotto, 2011.)

and reproduction rather than the alignment of perception with physical features of images or real-world sources (see previous discussion).

More specifically, in Figure I.9-3c the cumulative probability distributions would be expressed by $P(<L_T> \mid L_S)$, where $<L_T>$ represents the full range of the possible luminance values of a target patch that have occurred in conjunction with a given luminance value of the surround (L_S). The distribution $P(<L_T> \mid L_S)$ therefore provides an empirical scale that ranks *all* the target luminance values that have co-occurred with the surround in question, indicating the percentage of target patches having luminance values less than T and the percentage of target patches with luminance values greater than T. Therefore, for any particular target luminance (T) along the abscissa, the corresponding rank along the ordinate (T*) predicts the relative perception of lightness of the target. As shown in Figure I.9-3c, the higher T* ranks on this empirical scale, the lighter the perception of that particular target patch.

Similar arguments and examples could have just as well been drawn for the perceptual domains of color, geometry, and motion, the other major perceptual categories in vision (Howe & Purves, 2005; Purves & Lotto, 2011; Purves et al., 2014; Purves et al., 2004; Purves et al., 2011). From these considerations, it follows that calling any perception of lightness a "visual illusion" is incorrect; instead, such perceptions simply reflect the way the visual brain generates *all* subjective responses to luminance. In these terms, then, the distinction between veridical and illusory perception is false.

DESCRIBING WHAT WE SEE ON A WHOLLY EMPIRICAL BASIS

If the physical properties of objects in the world (or retinal images) are not represented in vision, then what is the proper way to describe what we see? Some indication is provided by sensory qualities in other modalities where the answer is more obvious. For example, it is clear that the sensory quality of pain is not inherent in the physical objects or conditions that cause it; rather, pain is a subjective quality that we have evolved because such perceptions promote survival. In an empirical framework, visual qualities bear a similar relationship to the world. Although we routinely attribute visual perceptual qualities to objects and environmental conditions, our experiences of lightness, brightness, color, geometry, and motion are likewise subjective qualities that promote useful behavior. Accordingly, it would be best to describe visual perceptions in terms similar to those used to describe pain, for which the concept of the "representation" does not make sense. Visual perceptions, like the perception of pain, do not stand for the actual properties of objects in the physical world, although the world is of course relevant to the way evolved visual systems determine stimuli and their frequency of occurrence (Purves et al., 2014).

This distinction between visual perception conceived in terms of the awareness of behaviorally useful qualities versus conceiving perception in terms of correct and incorrect representations of the physical world would be a theoretical point only were it not for the associated neurobiological implications. If vision does not represent the actual properties of objects and conditions in the world, then neither do its underlying anatomical and physiological mechanisms, which must therefore be thought of, examined, tested, and ultimately explained in empirical terms—a shift in thinking that leaves no place for the concept of "illusions."

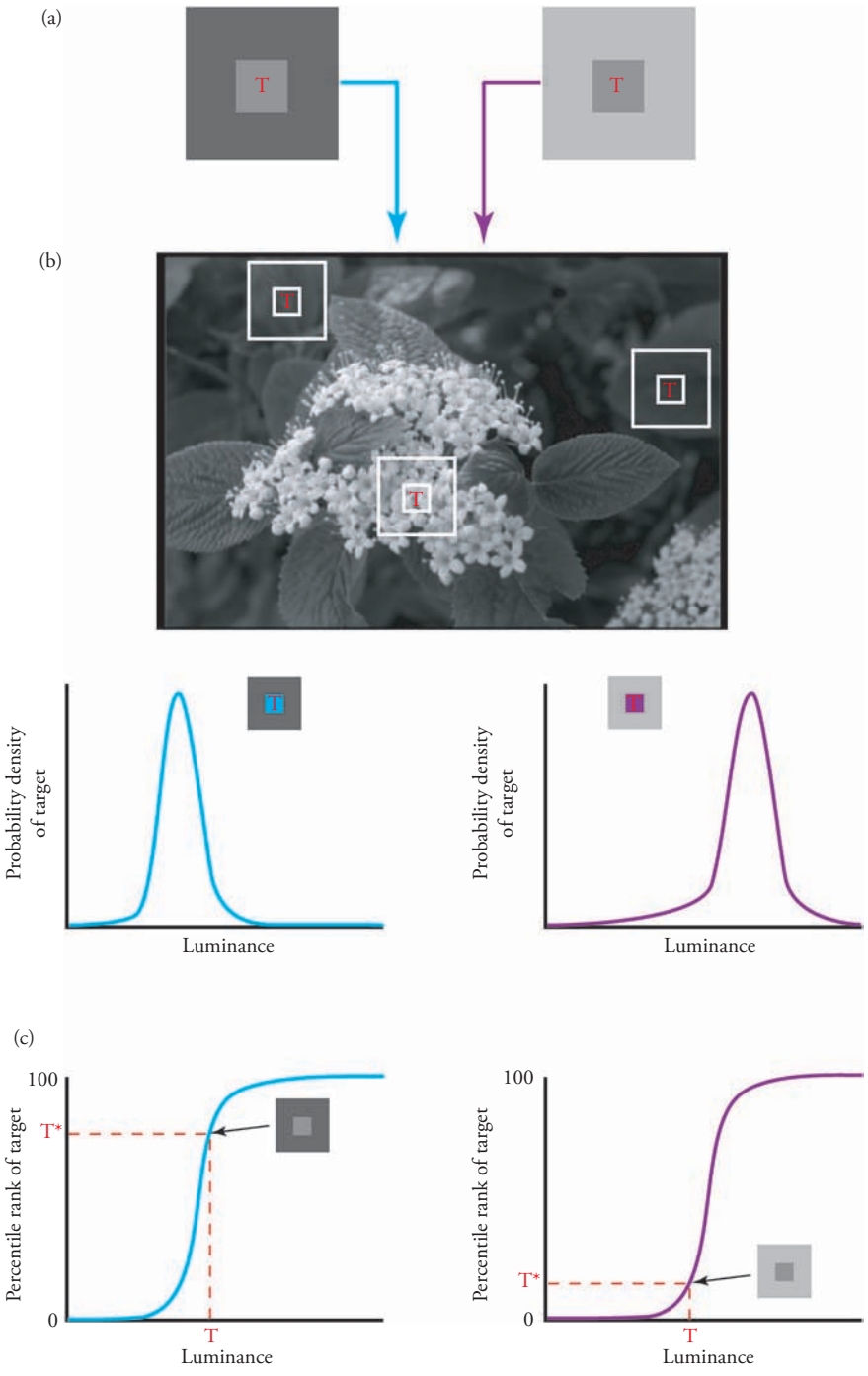


Figure I.9-3. Explaining the perception of lightness by the frequency of target and context luminance values in biologically determined stimuli. (Based on Yang & Purves, 2004.)

REFERENCES

Gelb, A. (1938). Die "Farbenkonstanz" der Sehdinge. In W. A. von Bethe, G. von Bergmann, G. Embden, & A. Ellinger (Eds.), *Handbuch der normalen und pathologischen Physiologie*. Band 12, 1. Hälfte Receptionsorgane II (pp. 594–678). Berlin: Springer. (English translation in W. D. Ellis [Ed.], *A source book of Gestalt psychology* [pp. 196–209]. New York: Harcourt Brace.) (Original work published 1929)

Gilchrist, A. L. (2009) *Seeing in black and white*. Oxford: Oxford University Press.

Howe, C. Q., & Purves, D. (2005). *Perceiving geometry: Geometrical illusions explained by natural scene statistics*. New York: Springer.

Hubel, D. H., & Wiesel, T. N. (2005). *Brain and visual perception*. New York: Oxford University Press.

Palmer, S (1999). *Vision science: From photons to phenomenology*. Cambridge, MA: MIT Press.

Purves, D., & Lotto, B. (2011). *Why we see what we do: An empirical theory of vision redux*. Sunderland, MA: Sinauer Associates.

Purves, D., Monson, B., Sundararajan, J., & Wojtach, W. T. (2014). How biological vision succeeds in the physical world. *Proceedings of the National Academy of Sciences USA*, *111*(13), 4750–4755.

Purves, D., Williams, M. S., Nundy, S., & Lotto, R. B. (2004). Perceiving the intensity of light. *Psychological Review*, *111*(1), 142–158.

Purves, D., Wojtach, W. T., & Lotto, R. B. (2011). Understanding vision in wholly empirical terms. *Proceedings of the National Academy of Sciences USA*, *108*(Suppl. 3), 15588–15595. doi:10.1073/pnas.1012178108

Purves, D., Morgenstern, Y., & Wojtach, W. T. (2015). Perception and reality: Why a wholly empirical paradigm is needed to understand vision. *Frontiers in Systems Neuroscience*. doi: 10.3389/fnsys.2015.00156.

Yang, Z., & Purves, D. (2004). The statistical structure of natural light patterns determines perceived light intensity. *Proceedings of the National Academy of Sciences USA*, *101*, 8745–8750.

Chapter 10
Where Have All the Illusions Gone?
A Critique of the Concept of Illusion

Brian Rogers

THE CONCEPT OF ILLUSION

Veridical and Illusory Perception

Illusions are usually described in terms of a discrepancy between some description of the physical reality and what we perceive. Veridical perception, on the other hand, refers to those situations when there is a correspondence between the physical reality and what we perceive. However, while there might be general agreement on what we perceive in particular circumstances, it is less clear how we should define physical reality. In 1922 Lukeish wrote: "any visual perception which does not harmonise with physical measurements may be termed an illusion" (p. 10). Barbara Gillam (1998) proposed that "the term illusion typically refers to a discrepancy between perceived reality and objective or physical reality" (p. 95). In 2004 Richard Gregory defined illusions as "discrepancies from truth" (p. 426) and in his last book, *Seeing Through Illusions*, he suggested that "illusions are departures from reality" (Gregory, 2009, p. 9). Nick Wade (2005) made the point that illusions require "a yardstick or reference relative to which (the illusion) can be assessed." Hence, in order to classify some aspect of our perception as illusory, we first need to identify and define the thing from which our perception is judged to be discrepant.

The "Stimulus"

The preceding discussion about how we should define the yardstick or objective reality from which our perception might be discrepant is closely related to the issue of how we should define the input or "stimulus" for the perceptual system. Philosophers often criticize psychologists for inadequately defining their terms, and this is very evident in the way we use the word "stimulus." The "stimuli" for individual receptors in the human eye are quanta of light (Gibson, 1960, 1968), but vision scientists and psychologists in general are more likely to use the term "stimulus" to refer to some higher-level description of the spatiotemporal patterning of wavelengths that reach the eye. However, this does not mean that there is agreement about what constitutes a stimulus. Some regard the optic array as the stimulus rather than the retinal image; others refer to the distinction between the "proximal" and the "distal" stimulus; and in other situations we may contrast the physical characteristics of the stimulus with the information the stimulus might contain. In the latter case, we often think that we are describing the physical characteristics of the stimulus when we are actually referring to the information the pattern of light contains. For example, we talk about a "ground plane surface," a "face," or "red fruit surrounded by green foliage"—and in doing so we are assuming (implicitly) that there are some common features in all ground plane surfaces, faces, or camouflaged fruit and, moreover, that there are mechanisms in our visual systems that can extract that particular information.

Optic Arrays and Retinal Images

Consider the first distinction between the spatiotemporal patterning of wavelengths at a vantage point—the "optic array"—and the "retinal image" (Frisby, 2009; Gillam, 2007; Rogers, 2007, 2009). Leaving aside for the moment the fact that our eyes are in constant motion so that the image on the retina changes in situations where the optic array remains the same, there is an important conceptual difference. The retinal image is a product of a particular type of visual system that has a chambered eye with a lens and an array of photoreceptors (Land & Nielsson, 2001). As a consequence, it cannot be regarded as a description of the input to visual systems in general. The concept of the optic array, on the other hand, is independent of the characteristics of any particular seeing machine—it is a description, in angular terms, of the patterns of light reaching a particular vantage point (Gibson, 1979).

To give an example: my students sometimes tell me that binocular stereopsis is based on the small differences or disparities between the two retinal images. In fact, binocular stereopsis exists as a source of information about the three-dimensional (3D) world because the optic arrays produced by a given 3D scene are different from two different *vantage points*. The retinal images in the two eyes are only different as a *consequence* of this fact.

Of course one needs to appreciate the characteristics of a given visual system in order to understand how information is extracted in that particular system, but the careful analysis of the spatiotemporal patterning of wavelengths in the optic array(s) is, in my view, logically as well as physically prior.

This particular example reveals a more subtle difference between retinal images and optic arrays that is not always made explicit. While both can be described in terms of the spatial patterning of wavelengths, either at a vantage point or over a projection surface, those who prefer to talk about optic arrays are more likely to describe the optic array in terms of the *information* the array contains (e.g., Gibson, 1968). In so doing, they go beyond a simple physical description of the stimulation. The importance of Gillam and Lawergren's (1983) and Mayhew and Longuet-Higgins's (1982) analyses of the vertical differences between the images reaching the two eyes (vertical disparities) lies not in the physical differences per se but rather in the fact that they provide *information* about the absolute distance to surfaces and the vergence angle of the eyes. The search for the information that might support any particular perceptual ability is what Marr (1982) referred to as the "computational theory" of vision.

The "Proximal Stimulus" as the Truth or Yardstick

Before we consider the concept of information in more detail, let us consider whether the "proximal stimulus" might be an appropriate basis for deciding whether a particular perception is illusory. The characteristics of the proximal stimulus (the retinal image) are both objective and a physical reality, and they could certainly be regarded as a yardstick or a reference. Moreover, they satisfy Gregory's (1994) description of the "truth" as "the world which is described by the Natural Sciences" (p. 93), since the characteristics of the proximal stimulus—its size, shape, and spectral content—can be measured objectively with rulers and other instruments. However, despite the truth and objective reality of the proximal stimulus, it is clearly not an appropriate basis for identifying and defining what is illusory. The fact that I see the shape of my desk as rectangular even though the image of the desk on my retina is trapezoidal in shape does not make my perception of the desk "illusory."

A similar argument applies to all our perceptual constancies. The size and shape of retinal images, the disparities between the two retinal images, and the particular wavelengths of light reaching the receptors may be discrepant from the characteristics we perceive, but this does not make our perception of the size, shape, depth, and color "illusory," One hundred and fifty years ago, Helmholtz (1910/1962) commented that "neither the size form and position of the real retina nor the distortions of the image projected on it matter at all . . ." (p. 167). In my view, it is of little consequence that we are given "tiny, distorted and upside-down images" (p. 7) (Gregory, 1966)—since image properties such as these are not necessarily the most appropriate or the most useful way to describe the input to our perceptual systems. Moreover, by emphasizing the lack of correspondence between retinal image characteristics and what we perceive, we run the risk of making the processes of perception sound more difficult, mysterious, or even miraculous.[1] This is evident in some of the popular perception textbooks and TV programs where the difficulties and mysteries of perception are often emphasized at the expense of pointing out the consistencies and the richness of the perceptual information available.

The "Distal Stimulus" as the Truth or Yardstick

So far we have rejected the proximal stimulus as being the appropriate "yardstick" or "truth" from which to identify what should be regarded as illusory, but what about the alternative—the "distal" stimulus—usually defined as what is "out there in the world"? This sounds more promising since there is typically a fair degree of correspondence between, for example, the color (reflectance property) of a surface and what we perceive. In other words, by using a distal stimulus description as the "truth," the constancies of our perception are no longer regarded as illusions. Extending the argument, it would not make sense to label perfect constancy as "veridical perception" and small departures from perfect constancy as "illusions." This would certainly be the case when failures of perfect constancy can be attributed to an insufficiency of the available information (Glennerster, Rogers, & Bradshaw, 1998; Holway & Boring, 1941). There are, however, some situations where the properties of the distal stimulus provide a misleading "truth" for deciding whether or not our perception is illusory.

The Ames Room

Consider the Ames room (Ames, 1951). Under ideal conditions, a properly constructed Ames room will create exactly the same pattern of light at the monocular peephole as a normal rectangular room (Runeson, 1988). In the first edition of *Eye and Brain* in 1966, Richard Gregory wrote, "it must look like a normal room if constructed according to strict perspective, and viewed from the right position, because the image it creates is the same as for an ordinary room" (p. 180). Hence, it makes no sense to ask the question "Why don't we see the room as trapezoidal?" because, if it is built with sufficient care, there is no visual information for its true, trapezoidal shape (Rogers, 2004).[2] The more appropriate question to ask is why a normal rectangular room looks rectangular, and this question leads us to consider the nature of perspective information as well as to investigate human observers' sensitivity to that information.[3]

Glennerster, Tcheang, Fitzgibbon, and Parker (2006) have taken a different view. They described the Ames room in the following way: "In that illusion, the two sides of the room have different scales but appear similar because observers fail to notice the gradual change of scale across the spatial extent of the room" (p. 431). While it is true that there is a gradual change of scale in the dimensions of an Ames room, the crucial point is that there is no information in the pattern of light reaching the eye from a properly constructed Ames room about that change of scale—there is nothing to "notice"! It follows that no seeing machine

could distinguish between the two situations, and therefore the demonstration of the Ames room by itself is essentially trivial.

In *Eye and Brain*, Gregory (1966) argued that the Ames room becomes interesting when people are added into the room because it shows that the "perceptual interpretation involves betting on the odds" (p. 180). I agree—adding people into the Ames room makes the demonstration more interesting because there are now two conflicting sources of information—the perspective information from the room indicating that it is rectangular and the size-of-a-familiar-object information suggesting that the two people should be at different distances (since their angular sizes are different). Note that the size-of-familiar-object information does not actually contradict the perspective information about the *shape* of the Ames room—only that the people should be at different distances.

However, it is also important to appreciate that putting these two sources of information into conflict has nothing to do with the fact that the Ames room is trapezoidal. The demonstration could have been done just as well in a normal, rectangular room by superimposing two different-sized images of the same person into the two corners (Fig. I.10-1). The Ames room was just a convenient way of manipulating the two sources of information before more sophisticated techniques became available. It should be added that situations like the Ames room where two or more sources of information (including tactile cues) are put into conflict can be informative because they allow us to investigate which source of information has the greater influence on what we see. It is not clear, however, that what we perceive should be labelled as "illusory," whatever the outcome.

Stereoscopes

The viewing of two flat pictures in a stereoscope provides a second example of a situation that should not be regarded as an illusion according to any definition. If the patterns of light reaching the two eyes are identical to those that would be produced when viewing a real 3D scene, then it is not surprising that what we see corresponds to what is seen in the real 3D scene—no seeing machine could ever tell the difference. My students sometimes tell me that the depth seen in a stereoscope is illusory because the two pictures (the distal stimuli) are actually flat. In order to convince them that it does not matter that the stereo pictures are flat, I ask them to consider the situation of viewing the world, not directly, but via a pair of small two-dimensional screens in a head-mounted display that receive their inputs from two cameras positioned directly in front of the screens. If the optic arrays at the two eyes are identical to those created when viewing the scene directly then, once again, no seeing machine could distinguish between the two situations and hence it makes no sense to talk about one situation giving "veridical" perception and the other "illusory." Of course this is not to say that stereoscopes are not very useful devices that allow us to manipulate binocular disparities and thereby learn about the way our visual system extracts disparity information. It does, however, make the important point that the crucial feature is the *information* provided by those binocular disparities rather



Figure I.10-1. "Ceci n'est pas une Ames room?" (Image by Brian Rogers.)

than the particular way the patterns of light are created—a point we return to later.

Structure-From-Motion Transformations

Structure-from-motion transformations provide a third example of situations that should not be regarded as illusory. If the transforming pattern of motion on a computer screen is identical to that produced by an actually rotating 3D structure, how could we see anything else? In the case of the kinetic depth effect (KDE; Wallach and O'Connell, 1953), a distinction is made between parallel and perspective projection of the displayed image. It is often claimed that perspective projection mimics viewing at close distances while parallel projection mimics viewing at infinity. In fact, the changes in the vertical height of features of an object rotating about a vertical axis are characteristic of objects that subtend a large visual angle, irrespective of whether that object is close to us or farther away (Braunstein, 1976; Rogers, 1993). Hence, displaying a KDE transformation under parallel projection that subtends a significant visual angle represents an impossible real-world situation. The perceptual consequence of this impossible real-world event—a figure that reverses both its perceived depth and direction of rotation is interesting, but once again it is not clear that the effect should be regarded as illusory. If there is no information about the direction of rotation, as would normally be provided by vertical perspective or occlusion (Braunstein, 1966; Braunstein, Anderson, Rouse, & Title, 1986), then both percepts (including their associated directions of rotation) are consistent with the available information. The same logic applies to other ambiguous figures such as the Necker cube—both "interpretations" are consistent with the stimulus and hence it does not make sense to talk about one or the other (or both) percept(s) as illusory.

A similar case can be made for the depth perceived in the random dot corrugations used by Rogers and Graham (1979) to investigate motion parallax. The distal stimulus was a random array of dots on a flat screen that was subject to a sinusoidal pattern of shearing motion linked to the position of the observer's head. Like the KDE, the patterns of relative motion are effectively identical to those that would be produced by the peaks and troughs of a real 3D surface during observer movement. As a consequence, it would not make sense to label the observer's percept as illusory in the experimental situation but veridical in the case of viewing a real corrugated surface.

Impossible Motion: Magnet-Like Slopes

There are many other examples of situations that should not be labeled as illusory on the grounds that what we perceive does not correspond to the physical reality of the situation that created the percept. Consider, for example, Kokichi Sugihara's (2010) striking entry in the 2010 "Best Illusion of the Year" contest. In his winning illusion—"Impossible Motion: Magnet-Like Slopes"—we perceive balls apparently rolling uphill along a set of upward-sloping runways (Fig. I.10-2a). In reality, the runways are actually downward sloping—as is revealed when the same scene is viewed from a different position (Fig. I.10-2b)—but they are constructed so as to provide perspective information indicating their upward slope from one particular viewing position. From this position, the pattern of light reaching our eyes is a copy—a facsimile—of the pattern of light that would be produced by actual, upward-sloping runways.

Like the Ames room, there is a mismatch between the physical reality—a downward-sloping runway or a trapezoidal room—and what we perceive. At the same time, there is, and *has to be*, a consistency between the information provided by the patterns of light reaching our eyes and what we perceive, and because of this it seems inappropriate to label what we perceive as illusory. Moreover, like the Ames room, viewing the (apparently) upward-sloping runways by themselves tells us nothing about the underlying perceptual processes since no seeing machine could tell the difference. In the case of Sugihara's (2010) "Impossible Motion" illusion, the cue conflict is not between different sources of perceptual information, as it is in the Ames room, but instead between the perspective information indicating that the runways are sloping upward and our cognitive knowledge that balls do not run uphill. Kokichi

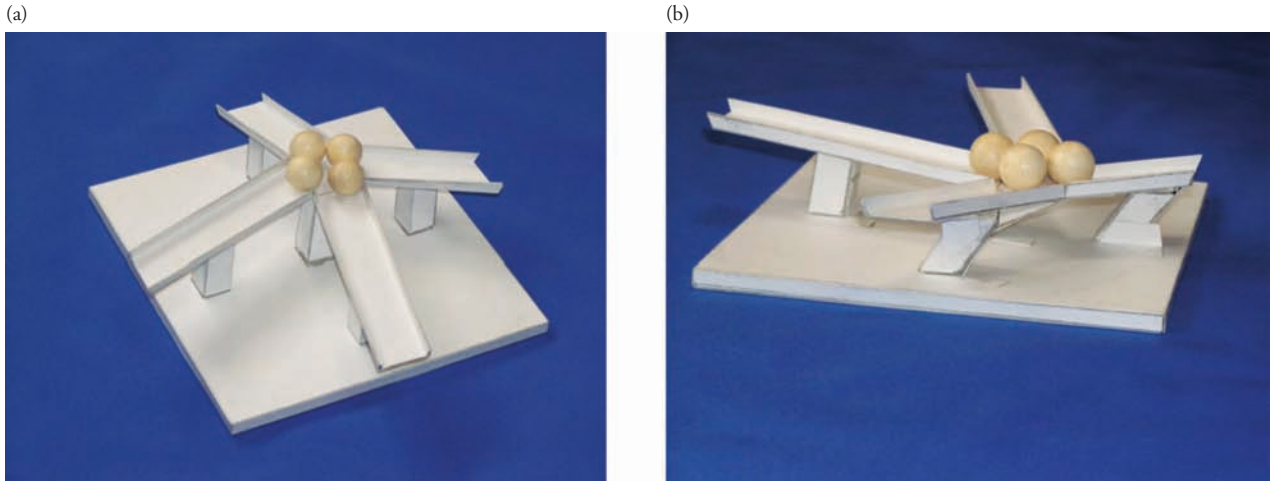


Figure I.10-2. The appearance of upward-sloping runways when seen from a particular vantage point (a). The reality of the physical situation that can be discovered from a different vantage point (b). (Image by Kokichi Sugihara.)

Sugihara's beautiful and amusing demonstration shows that the perceptual information wins out.

When Is an Illusion Not an Illusion? When It Is a Facsimile

Constructing either a stereogram or a transforming optic flow pattern on a computer screen should be seen, like the Ames room, as a convenient way to manipulate the *information* presented to the visual system, but the way that the particular pattern of light is created is surely totally irrelevant to our understanding of the perceptual system. Moreover, describing the input in terms of how the pattern of light is created (e.g., "*two-dimensional* stereo images" or a "*flat* computer screen" or the "*trapezoidal* shape of the Ames room") rather than the real-world situations they are mimicking has the potential to cause much confusion.

What these four scenarios have in common is that each represents a different way of presenting a pattern of light to the eyes that is essentially identical to the pattern of light that would be generated by a different real-world situation. The patterns of light are *facsimiles* ("projectively equivalent"; Howard & Rogers, 2002) and, as well as rejecting them as illusions, it is clear that facsimiles cannot tell us anything about the nature of our perceptual processes that could not have been learned from studying the situations they are mimicking.

Two important points come out of this discussion. First, it is irrelevant how a particular pattern of light reaching the eye (or eyes) is created, and second (and closely related), what is important for deciding whether a perception is illusory is not the description of the distal stimulus in purely physical terms but rather a description of the information the pattern of light contains.

Information Rather Than the Distal Stimulus

A possible way out of the problem of how to define the "truth" or "physical reality" would be to define an illusion as *a discrepancy between the information contained in the spatiotemporal patterning of light and what we perceive*. In the case of the Ames room (or indeed a normal rectangular room), we correctly perceive its rectangularity because there is (has to be) information, such as perspective and texture gradients, to specify its 3D shape. In the case of viewing images in a stereoscope (or the actual real-world scene), we perceive the depth because there is information in the pattern of disparities. By defining the "objective reality" or "truth" in terms of information available, rather than the physical characteristics of the distal stimulus, we get around the problem of facsimiles, but at the same time it raises the important question of how we should define "information" for our perceptual systems. Note that the concept of information is central to both Gibson's (1979) view of perception as a process of direct "information pickup" and David Marr's (1982) notion of a "computational theory."

A partial answer to the question of what constitutes "information" is provided by geometry and mathematics, and this is nicely illustrated using the vertical disparity example described earlier. Geometry shows that there is information about absolute distance and the vergence angle of the eyes in the pattern of vertical disparities between the two optic arrays (Gillam & Lawergren, 1983; Mayhew & Longuet-Higgins, 1982). Empirical evidence has since shown that vertical disparities are used by the human visual system (Rogers & Bradshaw, 1993).

However, the reason geometry and mathematics provide only a partial answer (at least at present) is that we need to be more precise about which particular aspect of the information is used by the visual system. In the case of the vertical disparity information used to determine the vergence angle of the eyes and absolute distance, is it the point disparities, orientation or curvature disparities, disparity gradients (and, if so, at which spatial scales), and so on? Similarly, it is not sufficient to say that our perception of a rectangular room is based on the perspective information present in the Ames room—is it the convergence of lines, gradients of size, foreshortening, texture density, or some combination of all four? The lack of precision in identifying what constitutes information is also apparent in the claim that there is *information in the patterns of wavelengths reflected off a number of surfaces to specify the color of a surface*. These become questions at Marr's (1982) "algorithmic" rather than "computational theory" level and the lack of specificity in identifying what constitutes "information" lies behind the frequent criticism that the Gibsonian concept of information is too hand-waving.

David Lee's (1980) insightful identification of the inverse relationship between the rate of dilation of the expanding optic array (or the retinal image) and the "time-to-contact" between an observer and an approaching object provides a second example of why geometry and mathematics provide only a partial answer to the question of what constitutes information. Until we know more precisely whether it is the rate of change of the size or texture density of the elements or the relative velocity of oppositely moving contours that is used by the human (or other visual systems), we cannot be certain as to what should be regarded as "the information for time-to-contact." In other words, not all *information-in-principle* turns out to be *information-in-practice*, as is evident in the case of the spatiotemporal patterning of infrared radiation reaching the human eye. Note that only empirical evidence, rather than geometry and mathematics, can provide the answer to this question.

Even though we do not know all the answers to the question of what constitutes information-in-practice for the perceptual system at present, it seems obvious that when we do reach that point, not only will we have identified and specified the relevant information, but we will also have described how our perceptual system works. In other words, a complete and correct definition of *information-in-practice* is also a description of how our perceptual system works, and a complete and correct description of how our perceptual system works is equivalent to a definition of what constitutes *information-in-practice*. In other words, "information" and "how our perceptual systems work" are two sides of the same coin (Gibson, 1979). Moreover, this discussion makes it clear that what constitutes information for one species—for example, the spatial patterns of infrared radiation for snakes—does not constitute information for other species. Information is species-specific just as perceptual mechanisms are species-specific (Land & Nielsson, 2001; Michaels & Carello, 1981).

Information and the Concept of Illusion

The mutuality of information and how our perceptual systems work (Gibson, 1979) has an important consequence for the concept of illusion. If what constitutes information for the visual system is defined with sufficient precision, it is clear that it does not make sense to define illusions in terms of a discrepancy between the information available and what we perceive because there can be no discrepancies and hence no illusions—there are only descriptions of how our perceptual systems work.

This conclusion is nicely illustrated using the example of metamerism. We do not regard our perception of color metamers (e.g., a mixture of red and green light that is seen as yellow) as an illusion but rather as just the way that our trichromatic vision system works. Note that since all perceived colors have metamers, *all* color perception would have to be regarded as illusory and not just metameric yellow, as Charles Wheatstone pointed out 150 years ago.[4] Likewise, we do not regard the fact that there are thresholds or other nonlinearities in the visual system as illusions—they are *just the way the system works*. The issue is essentially about terminology. It should not surprise us that, for most of the time, our perceptions provide us with more or less correct answers as to what is out there in the world, because our perceptual system would never have evolved if this had not been the case. However, there will always be limitations on perceptual performance that are a consequence of the particular ways in which the perceptual system works—for example, failures to perceive anything (absolute thresholds); failures to discriminate (including metamers); nonlinearities, and other systematic distortions of magnitude; and so on. At present, it seems that a subset of our perceptions are classified as "illusions" while others are just treated as *just the way the system works*. I am not convinced that there is a justifiable basis for making this distinction.

TEST CASES FOR THE CONCEPT OF ILLUSION: SIMULTANEOUS AND SUCCESSIVE ASSIMILATION AND CONTRAST EFFECTS

Many of the perceptual phenomena that are traditionally regarded as illusions can be classified under the headings of (a) simultaneous and (b) successive assimilation and contrast effects (Anstis, 1975). *Simultaneous* effects refer to perceptual distortions due to the (simultaneous) influence of neighboring stimuli along a particular sensory dimension and *successive* effects as the subsequent (successive) perceptual distortions as a result of prolonged stimulation along a particular sensory dimension. Examples of the former include Duncker's induced movement, depth, color, and tilt contrast; Mach Bands; and the Hermann-Hering grid. Examples of the latter include motion, depth, figural and tilt aftereffects, as well as luminance and color afterimages.

Simultaneous effects are often "explained" in terms of either lateral excitatory processes (assimilation effects) or inhibitory processes (contrast effects), while successive contrast effects are seen as a consequence of recurrent inhibitory processes (Anstis, 1975). Both processes can be involved as revealed by Anstis, Rogers, and Henry's (1976) "aftereffects from induced colour" and "induced colour from afterimages." "Explained" is in quotes because while excitatory and inhibitory interactions may be characteristics of a possible processing mechanism, those interactions tell us little about the *function* of such mechanisms in the visual system. The receptive fields of retinal ganglion cells have both excitatory and inhibitory influences, but this does not tell us anything about the function or role of those receptive fields: Is their purpose to detect features, identify zero crossings, analyze spatial frequencies, or reduce redundancy (Barlow, 1961)? Perhaps we should think about receptive fields simply as neural mechanisms that filter the input in particular ways, without attributing any explicit purpose.

Hence, an alternative way of characterizing the perceptual effects of assimilation and contrast processes would be to say that assimilation effects are a consequence of low-pass filtering and contrast effects the consequence of high-pass filtering. Returning to the question of what constitutes an illusion, it is interesting to note that the perceptual consequences of high-pass filtering are typically thought of as illusions, whereas the consequences of low-pass filtering are more often regarded as mere limitations of the perceptual system. Is there any justification for making such a distinction?

Simultaneous Assimilation Effects

There is a large class of illusory effects that have been attributed to lateral *excitatory* processes in the visual system, such that closely spaced stimuli in the physical world are no longer separately discriminated and what we perceive corresponds to some sort of averaged response. Some assimilation effects may be a consequence of the optical blurring of fine detail. If this is the case, it would not make sense to classify, for example, the inability to see fine detail at large distances as an illusion. In other cases, the spatial averaging occurs at scales larger than the retinal grain, and the assimilation effects are assumed to be a result of physiological rather than optical averaging processes—"neural blurring." However, like optical blurring, I assume that no one would wish to classify the poorer acuity that is characteristic of the peripheral retina as illusory. To take an example from binocular stereopsis, it has been shown that depth corrugations cannot be seen when the corrugation frequency is higher than about 4 cycles/deg (Rogers & Graham, 1982; Tyler, 1975). Once again, the failure to see corrugations with a frequency higher than 4 cycles/deg is not regarded as an illusion but rather as a consequence of the limited resolving power of the disparity mechanisms.

In 1975 Anstis and Rogers proposed that "reversed apparent motion" (first identified by Anstis, 1970) was not the result of a failure of the motion-detecting mechanisms per se but rather a simple consequence of spatial summation in the visual system. They showed that spatial summation would cause the positive-plus-displaced-negative intensity profiles to be displaced in a direction *opposite* to the displacement between the original positive and negative contours. Their conclusion—that reversed apparent motion is not a consequence of the motion mechanisms per se—was

strengthened by the discovery of similar effects in the disparity domain—reversed stereopsis—and reversed vernier alignment (Anstis & Rogers, 1975; Rogers & Anstis, 1975). If this explanation is correct, it is interesting to speculate as to why reversed apparent motion and reversed stereopsis are typically categorized as illusions while other effects of optical or neural blurring (low-pass filtering) are merely regarded as the consequence of the poor spatial resolution in the underlying mechanisms.

Simultaneous Contrast Effects

Simultaneous contrast effects or induced effects can be seen as the complement to the simultaneous assimilation effects just described. In a typical experimental setup, a small test stimulus is surrounded by an inducing field and the perceptual consequence is that our perception of the test stimulus is biased *away* from the characteristics of the inducing field. For example, surrounding a vertical test grating with a grating that is tilted slightly clockwise results in the vertical stripes of the test patch appearing to be tilted slightly counterclockwise—called simultaneous tilt contrast (Anstis, 1975). It seems quite likely that lateral inhibitory processes are involved in many of the simultaneous contrast effects, but does that make the effects themselves illusory? To answer this question, we need to return to the question of what constitutes the most appropriate description of the input from which our perception is discrepant. In most cases of simultaneous contrast, the "input" is usually defined in terms of the motion, depth, color, or orientation of the *local* test figure that is surrounded by neighboring stimuli that have a slightly different motion, depth, color, or orientation. But limiting the description of the input to the characteristics of the local test figure is a very restricted description. No one would regard the discrepancy between the wavelengths of light reflected from a single test surface and its perceived color as an illusion because we know that the reflected wavelengths from a single surface are confounded by the spectral composition of the illumination. The wavelengths of light reflected off a single surface cannot provide information about the color of that surface. The mistake is to treat the purely *local* stimulus—the light reflected off a single surface—as an appropriate description of the input.

In the case of motion perception, the motion of an image across our retinas cannot tell us about the absolute motion of an object in the world because image motion (or the lack of it) is confounded by the consequences of our own eye movements. Information about object motion is available, in principle, if eye movement signals as well as retinal image motion signals are available. However, given that we have rather poor information about the movements of our eyes (particularly when tracking an object), it would make good sense to rely instead on the fact that the surrounding world does not typically move. Hence, if there is relative motion between a small object and the large surrounding field, it is more than likely that it was the object that moved, rather than the surround. If this were the case, we would inevitably suffer from the "illusion" of induced motion. Note that there is no need to assume that high-level processes[5] or decisions are involved in "attributing" the relative motion to the smaller or surrounded object (Al-Haytham, 1083/1989; Rock, 1983). A simple mechanism that compared low-level motion signals would predict the same effect.

In depth perception, the disparity of a single object is confounded by the vergence state of the eyes, and there are other similar examples. In many of these cases, the visual system appears to have evolved the strategy[6] of extracting the relative (in spatial terms) rather than the local, absolute characteristics in order to deal with these confounding factors. In the case of the figural effects such as the Hering, Zöllner, and Orbison illusions, they are classified as illusions because of a similarly limited definition of the stimulus that is based on the orientation of *local* stimulus elements. In doing so, we are making the assumption that the visual system is capable of parsing the visual world in this way in order to extract perceptual information.

The Müller-Lyer illusion is regarded as an illusion because we perceive the lengths of the arrowhead shafts as different even though the shafts have the same physical length. However, once again we are assuming that the visual system is capable of measuring the lengths of just one part of the overall stimulus configuration (the shafts) and ignoring the remainder. Carrasco, Figueroa, and Willen (1986) have suggested a possible explanation of the illusion based on the fact that the outputs of large, elongated receptive fields (low spatial frequency filters) would be different for the Müller-Lyer figure with outgoing fins compared with those for the ingoing fins version. Perhaps if we had a better understanding of what the visual system is trying to do—that is, extract information about length given certain confounding factors—we would no longer regard these effects as illusions (Morgan, Hole, & Glennerster, 1990). This suggests that many of the things that we label as illusions at present are merely those aspects of our perception that we do not fully understand. Once we understand what is going on, an effect that has been labeled previously as an "illusion"—for example, a metameric color match—becomes no more than an illustration of how the perceptual system works.

Successive Assimilation Effects

Successive assimilation effects occur when the visual system averages stimuli that change rapidly over time. The examples that come to mind include the fact that flicker is not seen when the flicker rate is above the critical flicker fusion frequency and the additive color mixing effects seen with rapidly alternating colors. Like the simultaneous assimilation effects described earlier, these successive assimilation effects would be regarded by most researchers as a consequence of the poor temporal resolution in the underlying mechanisms rather than illusions.

Successive Contrast Effects

In 1979 John Frisby wrote: "One large group of illusions . . . relies upon the fact that certain curious illusory phenomena are experienced after a period of prolonged or intense constant stimulation. Such illusions are called after-effects

and occur widely in sensory systems" (p. 89). In the case of one of the oldest effects in the history of visual science—the motion aftereffect—motion in the opposite direction is seen when the previously moving pattern stops moving. Surely this should be classified as an illusion. After all, we perceive motion in a stimulus that is stationary. There is certainly a discrepancy between the *momentary* stimulus (a stationary pattern) and what we perceive (a motion aftereffect). However, like the simultaneous contrast effects described earlier, that classification could be a consequence of an overly limited definition of the input. Just as a description of the input in terms of a *local* stimulus feature in space could be considered inappropriate for understanding and explaining simultaneous contrast effects, perhaps the description of the input as a *momentary* stimulus feature *in time* might also be considered inappropriate for thinking about successive contrast effects.

Once again, there seems to be an inconsistency in the way we treat assimilation and contrast effects. In general, we are quite willing to regard assimilation processes over time as an unfortunate limitation of our perceptual system, whereas successive contrast processes—aftereffects—are regarded as illusions.

ALTERNATIVE EXPLANATIONS FOR OTHER FAMILIAR ILLUSIONS

Filling-In

We are all familiar with the fact that we are not aware of the blind spots in the peripheral retina where there are no receptors. Moreover, a contour or pattern that extends across the blind-spot region of the retina does not appear to show a break. The usual explanation of this phenomenon is "filling-in," and it is regarded as an illusion because what we see—a homogeneous field or continuous pattern—does not correspond to activity (or lack of activity) in the retinal receptors (Ramachandran & Gregory, 1991). The mistake in this case seems to be the assumption that the visual system is trying to create some sort of internal picture that needs to be "filled-in." If instead we think of perception as a process of extracting information, then it is not surprising that we do not see our blind spots because there are no mechanisms that could extract any change of luminance (or any change of pattern) occurring at the boundaries of the blind spot. In any visual system, it is not sufficient simply to re-present the values of the luminances in different parts of the visual field at a subsequent processing stage (i.e., making a copy of what was present in the retinal image)—the visual system needs mechanisms to make those differences *explicit,* as Marr (1982) stressed over 30 years ago. Ganglion cell receptive fields and zero-crossing calculations make spatial luminance changes *explicit* everywhere in the visual field apart from at the blind-spot boundaries, and hence there is no explicit information from those regions to signal that the luminance or the pattern has changed.

To take the preceding argument to its limit, I think most readers would agree that there is no need to postulate a "filling-in" process to account for the fact that we do not perceive a surface with unchanging luminance as punctate because there are gaps between the individual cones or rods in the retina. Once again, there are no mechanisms (and there could be no mechanisms) to compare the activity of an individual receptor with the hypothetical activity from the region between that receptor and its neighbors and hence there is no information that the surface is anything but unchanging and continuous.

Hollow Masks

The familiar hollow mask demonstration, like many other ambiguous figures such as the Necker cube and Jastrow's duck-rabbit, relies on the fact that a rather impoverished stimulus to the visual system can be produced by two different real-world scenarios. The shading patterns created by a hollow mask and a normal (protruding) mask are extremely similar (Fig. I.10-3), and, even if the observer is looking directly at the hollow mask (rather than a picture of it), the disparity information about its 3D shape from the shallow luminance gradients is very weak or nonexistent (Arndt, Mallot, & Bülthoff, 1995). As a consequence, what we perceive—a normal (convex) face—is actually fully consistent with the only information that is available, and, like the Ames room, it is difficult to see why the effect should be labeled as an illusion other than the fact that the mask itself is hollow. Once again, we make the mistake of



Figure I.10-3. Which is a photo of the hollow mask and which is a photo of the protruding mask? (Image by Brian Rogers.)

attributing too much importance to the physical reality of the stimulus out there—a *hollow* mask—rather than analyzing the information that the stimulus provides, which, in this case, is for *either* a convex *or* a concave structure. If we alter the available information by scattering random dots over the surface of the mask to introduce significant disparities, what we perceive is a hollow mask, and this is consistent with both the disparity and the shading information (Georgeson, 1979).

The importance of considering the information available rather than the physical reality of the stimulus out there is highlighted when one considers the perception of "virtual" hollow masks (Rogers, 2010a, 2013). The "virtual" masks that we created allowed us to manipulate the particular shading patterns that are created by both normal and hollow masks, and we were also able to present different hollow-mask images to the two eyes and thus independently manipulate the disparities present. In addition, we were able to manipulate the different patterns of parallax motion created by our "virtual" normal and hollow masks by monitoring the observer's head position during side-to-side head movements. As a consequence, we were able to manipulate the *information* available in these situations (and monitor the resultant perceptual consequences) rather than thinking in terms of the *physical reality* of the stimulus.

Note that the hollow-mask demonstration differs from other ambiguous figures like the Necker cube in that we only ever experience one of the two possible percepts or "interpretations." We rarely perceive the mask as hollow even though such a percept is also consistent with the available information. Why should this be? Hill and Johnston (2007) presented evidence to suggest that there is an inherent bias in the visual system in favor of a convexity "interpretation" because of the statistics of real-world scenes—that is, there are more convex surfaces in the world than concave surfaces. Regardless of whether this is correct, it is important to note that as soon as information such as binocular disparities is added to what was initially a rather impoverished display, the ambiguity goes away. This raises the important question of whether it is wise to generalize about perceptual processes from the evidence gained from our perceptions under impoverished stimulus conditions where much of the information normally available is absent.

Reverspectives

Patrick Hughes's "Reverspectives" (Fig. I.10-4) have amused and intrigued both artists and visual scientists by creating an illusion of reversed depth and reversed apparent motion in the 3D pyramids and ridges that make up the artwork (Wade & Hughes, 1999; Slyce & Hughes, 2005). The parts of the artwork that are physically closest to the observer (the tops of the pyramids) are typically seen as farther away and hence the label—"Reverspectives." However, it should be pointed out that when a reverspective is viewed



Figure I.10-4. A "reverspective" artwork. From this oblique angle, readers can see that the images of the houses are printed on the sides of pyramids that protrude out of the background. On the other hand, those perspective images provide information that the houses are receding into the distance. (Image by Brian Rogers.)

monocularly by a *stationary* observer, there is little or no information about the real depth[7] of the artwork, and hence it is not surprising that what we see is consistent with the only available information (i.e., the perspective information presented on the faces of the pyramids and ridges). In this respect, reverspectives are very similar to the Ames room situation, and the fact that our perception is inconsistent with the physical depth in either the Ames room or a reverspective does not seem to be particularly important or to justify the label of an "illusion."

From the visual scientist's point of view, reverspectives become interesting when they are viewed with both eyes because this creates a conflict between the perspective information that the rows of houses are receding into the background and the disparity information that the pyramids and wedges are protruding out of the background. Under these conditions, observers typically report that the depth remains consistent with the perspective information, until the observer is less than a meter away (Papathomas, 2007; Rogers & Gyani, 2010). With binocular viewing of a reverspective, we have created a cue conflict, and it is not clear that we should regard our perception as illusory: either we see the depth as protruding—and this is consistent with the disparity information—or we see the depth as receding—and this is consistent with the perspective information. Moreover, is there any good reason to elevate the status of disparity information over the perspective information (note that we refer to them as *rever*spectives) beyond the fact that this was how the particular display was created? This point becomes particularly apposite when we construct "virtual" versions of reverspectives in which the disparity and perspective information is independently manipulated and the question of what the "real" depth is does not arise (Rogers, 2010a, 2013).

However, anyone who has viewed a reverspective artwork will have noticed that the "reversed depth" is enhanced when the viewer makes side-to-side or up-and-down head movements, and it is this property that makes reverspectives special and different from the Ames room. Observer movements create motion parallax between the parts of the artwork at different distances. Some authors have argued that the parallax information, like the binocular disparities, is incompatible with the perception of "reversed" depth (Cook, Hayashi, Amemiya, Suzuki, & Leumann, 2002, Papathomas, 2007). If this were the case, reverspectives would still represent a cue conflict situation, rather than an illusion, and we would be obliged to conclude that the perspective information wins out against the combined effects of the perspective and parallax information.

Rogers and Gyani (2009) have suggested an alternative interpretation. They proposed that the assumed incompatibility of the parallax information with our perception of the reversed depth rests on an assumption of that objects remain *stationary* during head movement. On this assumption, the motion parallax transformations are seen as being consistent only with the physical situation—a series of protruding pyramids and cones. However, if we abandon the assumption of object stationarity (Rogers, 2012), it becomes clear that the parallax transformations created by the pyramids and cones of a reverspective are also perfectly compatible with the reversed depth signaled by perspective. In this case, the "reversed" depth of the receding 3D structure would have to rotate concomitantly and in the same direction as the observer's movements, and, indeed, this is what we see. Although such a scenario is very unlikely to ever occur in real life, it does not appear to be vetoed by the visual system (Rogers & Collett, 1989).

Given this analysis, should we regard the perceived rotation of reverspectives as "illusory"? Once again, while the rotation is inconsistent with the actual (stationary) physical structure in front of the observer, it is quite consistent with the available parallax information. The only remaining inconsistency is with the disparity-specified depth. From a visual scientist's perspective, reverspectives raise interesting questions about the nature of motion parallax information and the concept of illusion, but this should not detract from the delight and amusement of these amazing artworks.

The Ponzo Illusion

In the Ponzo illusion, the two horizontal lines of equal length are seen as unequal when the lines are flanked by a pair of converging lines (Fig. I.10-5a). Many authors have pointed out that if we were viewing a real scene consisting of a railway tracks receding into the distance, the rails and the sleepers (ties) would create a retinal image that is similar to the abstract pattern of lines in the Ponzo illusion (Fig. I.10-5b). However, when viewing the real scene, the perceived longer length of the higher (more distant) line would be regarded as a correct rather than an illusory percept and a result of size constancy scaling.

What is the difference between the Ponzo illusion and the viewing of a real 3D scene that leads to the classification of our perception as illusory in the first case and as correct in the second? To try to answer this question, it may be helpful to imagine a continuum between the image created by the real-world scene (Fig. I.10-5b) and the impoverished Ponzo figure on a sheet of paper (Fig. I.10-5a). The first step on that continuum would be to remove the disparities from the real-world scene and replace them with the disparities of a flat, two-dimensional surface (i.e., a picture of the scene). Next, the perspective information of the real-world scene could be degraded by removing the texture gradients of the surroundings. This could be followed by the substitution of the rails (that have finite and changing angular widths) with a pair of thin lines that converge. Finally, we could reduce the number of sleepers to two before eliminating the thickness difference of the remaining two sleepers. Using this hypothetical continuum as a basis, would it make sense to ask at which point our perception should be labeled as changing from veridical to illusory?

One conclusion that could be drawn from this example it that if we take away more and more information, our perception suffers, and hence all we have learned from the Ponzo illusion is that. Moreover, we might also want to conclude that there are significant dangers in inferring how the perceptual system works in normal, information-rich situations from impoverished situations where some or most of the information normally used by the visual system is eliminated.

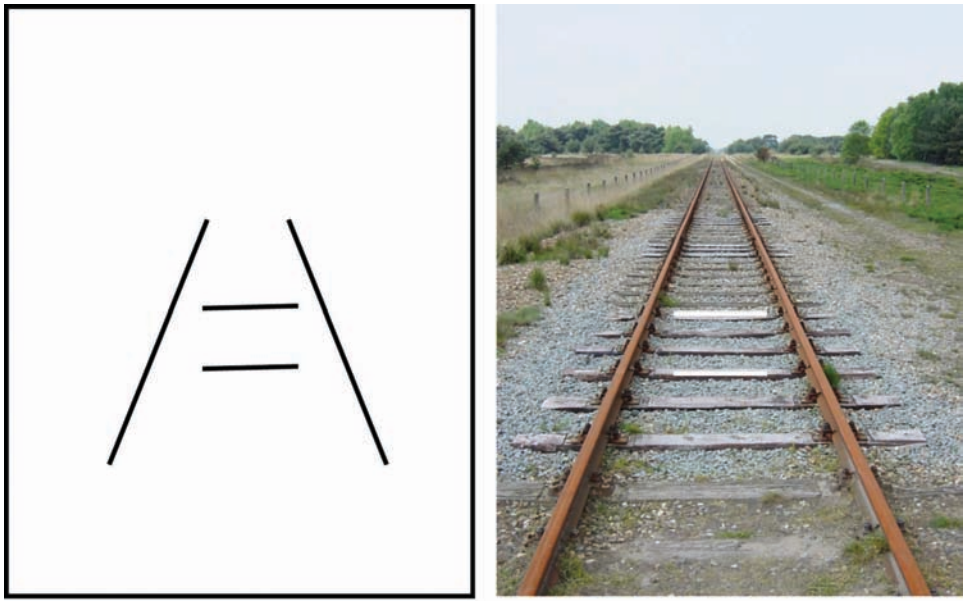


Figure I.10-5. The Ponzo illusion (a) and a photograph of a real scene that creates a similar retinal image (b). Note that there is a second, recently discovered "illusion": the angle between the two rails in the photo of the real scene (b) appears to be much smaller than the angle between the two inclined lines of the Ponzo illusion (a) (Osa et al., 2011). (Image by Brian Rogers.)

Sensory-Motor Adaptation

After sufficient exposure to displacing prisms that create a discrepancy between the seen and felt position of a target, we "adapt," and when the prisms are removed we point incorrectly to a visual target (Held, 1965). Should the consequences of sensory-motor adaptation after wearing displacing prisms be regarded as illusory? There is good evidence that the "felt position" of the hand and the eye are altered as a result of the adaptation rather than vision itself (Harris, 1965). Moreover, Ian Howard (1968) has shown that these adaptation effects can occur even when the observer is unaware of the displacing effect of the prisms, using a technique he called "prismatic shaping."

The winning entry in the 2012 "Best Illusion of the Year" contest was Newport, Gilpin, and Preston's "Disappearing Hand Trick," in which participants showed surprise and astonishment that they did not see their hand in the place where they felt it to be after a period of prismatic shaping. However, should we regard the consequences of wearing displacing prisms as an illusion rather than the manifestation of an adaptive process? Consider the process of dark adaptation. It is unlikely that anyone would regard our failure to see details of the surroundings when we first enter a dark room after being in bright sunshine as an illusion. Newport et al.'s delightful "Disappearing Hand Trick" makes the point that what we may classify as an illusion is not always based on some formal definition of a mismatch between what we perceive and the objective reality but instead is used as a label for those situations that surprise, amuse, or intrigue us most.

CONCLUSIONS

Two questions were raised at the start of this chapter. First, how should we identify the objective reality or truth from which our illusory perception is discrepant? Second, is there a satisfactory definition of an illusion that allows us to distinguish between those aspects of perception that are illusory and those that are veridical? The Ames room, the viewing of images in a stereoscope, and many other situations make it clear that the *distal* stimulus—what is out there in the world—does not always provide an appropriate basis for defining the objective reality. Second, the existence of projectively equivalent *facsimiles* such as the Ames room make the point that it does not matter how the particular patterns of light reaching the eyes are created—what matters is the *information* carried by those patterns of light. If we accept these two premises, then we might be tempted to conclude that a more appropriate use of the term "illusion" is for situations where our perceptions are discrepant from the available information.

However, such a definition raises the question of how we should identify the information that underlies a particular perceptual ability. Examples such as vertical disparities and time-to-contact show how geometry and mathematics can help us to identify and define what constitutes *information-in-principle*. However, before we can be sure what constitutes the *information-in-practice* that underpins a particular perceptual ability, we need to be able to specify much more precisely this thing we call information. Once this is done, we discover that the precise description of the *information-in-practice* becomes equivalent to a description of how the perceptual system works, and, as a consequence, there can be no illusions—only descriptions of how our perceptual systems work.

Does that make the study of the things that we loosely call "illusions" irrelevant? Not at all. The things that we call "illusions" should be treated like thresholds, failures to discriminate, metamers, nonlinearities, cue conflicts, and systematic distortions of magnitude—they all contain "visual truth" (Purkinje, 1825)—that is, they reveal something about how our perceptual systems work. In doing so, however, we might also have to acknowledge that there may be no meaningful way to distinguish between those perceptions that should be classified as "veridical" and those that should be classified as "illusory."

ACKNOWLEDGMENTS

An earlier article containing some of the ideas in this chapter was published as an editorial in *Perception* (Rogers, 2010b).

NOTES

1. Gregory (1966) went on to say: "We are given tiny, distorted upside-down images in the eyes, and we see separate solid objects in surrounding space. From the patterns of stimulation on the retinas we perceive the world of objects and this is nothing short of a miracle" (p. 7).
2. The far corners of an Ames room may require slightly different accommodations.
3. The same arguments apply with respect to the notion of "equivalent configurations": that there can be an infinite number of real-world configurations that produce the same pattern of light at the eye (Ittelson, 1960). While there can be no disagreement about the existence of "equivalent configurations," this idea has been used as evidence that perspective and other sources of 3D information are inherently ambiguous, but the argument is surely flawed. The ambiguity of a depth cue (or the lack of ambiguity) depends on (a) how the source of information is defined and (b) the characteristics of the particular mechanism used to extract the information, not on how the pattern of light was created.
4. Note that we are not talking about the qualia of the colors we perceive but our ability to identify and discriminate different shades.
5. Al-Haytham (1989) considered that induced motion was an "error of inference" in perception.
6. The use of the word "strategy" is not meant to imply any high-level process involving choice by the perceiver.
7. The front and back surfaces of a reverspective may require slightly different accommodation.

REFERENCES

Al-Haytham, I. (1989). Book of optics. In A. I. Sabra (Trans.), *The optics of Ibn Al-Haytham*. London: Warburg Institute. (Original work published 1083)

Ames, A. (1951). Visual perception and the rotating trapezoidal window. *Psychological Monographs, 65*, 1–31.

Anstis, S. M. (1970). Phi movement as a subtraction process. *Vision Research, 10*, 1411–1430.

Anstis, S. M. (1975). What does visual perception tell us about visual coding? In C. Blakemore & M. S. Gazzaniga (Eds.), *Handbook of psychobiology* (pp. 269–323). New York: Academic Press.

Anstis, S. M., & Rogers, B. J. (1975). Illusory reversal of visual depth and movement during changes of contrast. *Vision Research, 15*, 957–961.

Anstis, S. M., Rogers, B. J., & Henry, J. (1976). Interactions between simultaneous contrast and coloured afterimages. *Vision Research*, *18*, 899–911.

Arndt, P. A., Mallot, H. A., & Bülthoff, H. H. (1995). Human stereovision without localised image features. *Biological Cybernetics*, *72*, 279–293.

Barlow, H. B. (1961). Possible principles underlying the transformations of sensory messages. In W. A. Rosenblith (Ed.), *Sensory communication* (pp. 217–234). Cambridge, MA: MIT Press.

Braunstein, M. L. (1966). Sensitivity of the observer to transformations of the visual field. *Journal of Experimental Psychology*, *72*, 683–689.

Braunstein, M. L. (1976). *Depth perception through motion*. London: Academic Press.

Braunstein, M. L., Anderson, G. J., Rouse, M. W., & Title, J. S. (1986). Recovering viewer-centered depth from disparity, occlusion and velocity gradients. *Perception & Psychophysics, 40*, 216–224.

Carrasco, M., Figueroa, J. G., & Willen, J. D. (1986). A test of the spatial-frequency explanation of the Müller-Lyer illusion. *Perception, 15*(5), 553–562.

Cook, N. D., Hayashi, T., Amemiya, T., Suzuki, K., & Leumann, L. (2002). Effects of visual-field inversions on the reverse-perspective illusion. *Perception, 31*, 1147–1151.

Frisby, J. (1979). *Seeing: Illusion, mind and brain*. Oxford: Oxford University Press.

Frisby, J. (2009). Optic arrays and retinal images. *Perception*, *38*, 1–4.

Georgeson, M. (1979). Random-dot stereograms of real objects: observations on stereo faces and moulds. *Perception*, *8*, 585–588.

Gibson, J. J. (1950). *The perception of the visual world*. Boston: Houghton Mifflin.

Gibson, J. J. (1960). The concept of the stimulus in psychology. *American Psychologist*, *16*, 694–703.

Gibson, J. J. (1968). *The senses considered as perceptual systems*. London: Allen & Unwin.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Gillam, B. (1998). Illusions at century's end. In In J. Hochberg (Ed.), *Perception and cognition at century's end* (pp 95–136). San Diego, CA: Academic Press.

Gillam, B. (2007). Stereopsis and motion parallax. *Perception, 36*, 953–954.

Gillam, B., & Lawergren, B. (1983). The induced effect, vertical disparity, and stereoscopic theory. *Perception & Psychophysics, 34*, 121–130.

Glennerster, A., Rogers, B. J., & Bradshaw, M. F. (1998). Cues to viewing distance for stereoscopic depth constancy. *Perception, 27*, 1357–1365.

Glennerster, A., Tcheang, L., Fitzgibbon, A. W., & Parker, A. J. (2006). Humans ignore motion and stereo cues in favour of a fictional stable world. *Current Biology*, *16*(4), 428–432.

Gregory, R. L. (1966). *Eye and brain*. London: Weidenfeld & Nicolson.

Gregory, R. L. (1994). The unnatural science of illusions. *Royal Institution Proceedings*, *64*, 93–110.

Gregory, R. L. (2004). Illusions. In R. L. Gregory (Ed.), *The Oxford companion to the mind* (pp. 426–443). Oxford: Oxford University Press.

Gregory, R. L. (2009). *Seeing through illusions*. Oxford: Oxford University Press.

Harris, C. S. (1965). Perceptual adaptation to inverted, reversed, and displaced vision. *Psychological Review, 72*, 419–444.

Helmholtz, H. von. (1962). *Treatise on physiological optics*. New York: Dover. (Original work published 1910)

Held, R. (1965). Plasticity in sensory-motor systems. *Scientific American, 213*, 84–94.

Hill, H., & Johnston, A. (2007). The hollow-face illusion: Object-specific knowledge, general assumptions or properties of the stimulus? *Perception*, *36*, 199–223.

Holway, A. F., & Boring, E. G. (1941). Determinants of apparent visual size with distance variant. *American Journal of Psychology*, *53*, 21–37.

Howard, I. P. (1968). Displacing the optical array. In S. J. Freedman (Ed.), *The neuropsychology of spatially oriented behavior* (pp. 19–36). Homewood, IL: Dorsey Press.

Howard, I. P., & Rogers, B. J. (2002). *Seeing in depth*. Toronto: Iporteus.

Land, E. H., & McCann, J. J. (1971). Lightness and retinex theory. *Journal of Optical Society of America*, *61*, 1–11.

Land, M. E., & Nielsson, D.-E. (2001). *Animal eyes*. Oxford: Oxford University Press.

Lee, D. N. (1980). The optic flow-field: The foundation of vision. *Philosophical Transactions of the Royal Society of London, B*, *290*, 169–179.

Lukeish, M. (1965). *Visual illusions*. New York: Dover.

Marr, D. (1982). *Vision*. San Francisco, CA: Freeman.

Mayhew, J. E. W., & Longuet-Higgins, H. C. (1982). A computational model of binocular depth perception. *Nature, 297*, 376–378.

Michaels, C. F., & Carello, C. (1981). *Direct perception*. Englewood Cliffs, NJ: Prentice Hall.

Morgan, M. J., Hole, G. J., & Glennerster, A. (1990). Biases and sensitivities in geometrical illusions. *Vision Research*, *30*, 1793–1810.

Newport, R., Gilpin, H., & Thompson, R. (2012). *The disappearing hand trick*. Winner Best illusion of the year contest. https://youtu.be/Qcm0_QPhLe4

Osa, A., Nagata, K., Honda, Y., Ichikawa, M., Matsuda, K., & Miike, H. (2011). Angle illusion in a straight road. *Perception*, *40*, 1350–1356.

Papathomas, T. V. (2007). Art pieces that "move" in our minds—An explanation of illusory motion based on depth reversal. *Spatial Vision*, *21*, 79–95.

Purkinje, J. (1825). *Beobachtungen und Versuche zur Physiologie der Sinne*. Vol. *2*. Prague: Calve.

Ramachandran, V. S., & Gregory, R. L. (1991). Perceptual filling in of artificially induced scotomas in human vision. *Nature*, *350*(6320), 699–702.

Rock, I. (1983). *The logic of perception*. Cambridge, MA: MIT Press.

Rogers, B. J. (1993). Motion parallax and other dynamic cues for depth in humans. In F. A. Miles & J. Wallman (Eds.), *Visual perception and its role in the stabilization of gaze* (pp. 119–137). Amsterdam: Elsevier.

Rogers, B. J. (2004). Stereopsis. In R. L. Gregory (Ed.), *The Oxford companion to the mind* (pp. 878–881). Oxford: Oxford University Press.

Rogers, B. J. (2007). Optic arrays and celestial spheres. *Perception*, *36*, 1269–1273.

Rogers, B. J. (2009). "Three cheers for retinal images": A reply. *Perception*, *38*, 159–161.

Rogers, B. J. (2010a). Interactions between disparity, parallax and perspective: Linking "reverspectives," hollow masks and the apparent motion seen in random dot stereograms. *Journal of Vision, 10*, 55.

Rogers, B. J. (2010b). Stimuli, information and the concept of illusion. *Perception*, *39*, 285–288.

Rogers, B. J. (2012). Motion parallax, pursuit eye movements and the assumption of stationarity. *Journal of Vision, 12*(9), 251.

Rogers, B. J. (2013). *Virtual reverspectives*. Demonstration presented at the European Conference on Visual Perception.

Rogers, B. J., & Anstis, S. M. (1975). Reversed depth from positive and negative stereograms. *Perception, 4*, 193–201.

Rogers, B. J., & Bradshaw, M. F. (1993). Vertical disparities, differential perspective and binocular stereopsis. *Nature, 361*, 253–255.

Rogers, B. J., & Collett, T. (1989). The appearance of surfaces specified by motion parallax and binocular disparity. *The Quarterly Journal of Experimental Psychology*, *41*(4), 697–717.

Rogers, B. J., & Graham, M. E. (1979). Motion parallax as an independent cue for depth perception. *Perception*, *8*, 125–134.

Rogers, B. J., & Graham, M. E. (1982). Similarities between motion parallax and stereopsis in human depth perception. *Vision Research*, *22*, 261–270.

Rogers, B. J., & Gyani, A. (2010). Binocular disparities, motion parallax and geometric perspective in Patrick Hughes' "Reverspectives": Theoretical analysis and empirical findings. *Perception, 39*, 330–348.

Runeson, S. (1988). The distorted room illusion, equivalent configurations, and the specificity of static optic arrays. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 295–304.

Slyce, J., & Hughes, P. (2005). *Perverspective*. London: Flowers.

Sugihara, K. (2010). *Impossible motion: Magnet-like slopes*. Winner: Best illusion of the year contest. http://illusionoftheyear.com/2010/impossible-motion-magnet-like-slopes/.

Tyler, C. W. (1975). Spatial organisation of binocular disparity sensitivity. *Vision Research*, *15*, 583–590.

Wade, N. J. (2005). *Perception and illusion: Historical perspectives*. New York: Springer.

Wade, N. J., & Hughes, P. (1999.) Fooling the eyes: Tromp l'oeil and reverse perspective. *Perception*, *28*, 1115–1119.

Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology, 45*(4), 205–217.

PART II
Geometrical

Chapter 11
Weighted Positional Averaging in the Illusions of the Müller-Lyer Type

Aleksandr Bulatov

INTRODUCTION

Each of our sensory channels is adapted to encode though a fairly wide but still constrained range of information coming from the environment. Ordinarily, our senses are quite accurate, but in some circumstances the constraints lead to significant discrepancy (i.e., misperception or illusion) between the current sensory signaling and our knowledge concerning physical properties of the perceived objects. The purpose of this chapter is to discuss the relatively narrow class of so-called geometric-optical illusions of the Müller-Lyer type that affect judgments of spatial extent (or length) of images.

The most famous example of these illusions demonstrates the Müller-Lyer figures, named so after their discoverer, the German psychiatrist F. C. Müller-Lyer. Two equal lines (shafts) terminating in intersections of inward- and outward-pointing wings appear to be of different lengths (Fig. II.11-1a), and this effect may exceed 25% of the actual shaft length. Such strong illusion evoked by simple line drawings has attracted great interest from researchers, which has led to the appearance of innumerable modifications of stimuli. Müller-Lyer (1889) himself has proposed about a dozen patterns comprising the drawings of acute and obtuse angles in different spatial arrangement. Brentano (1892) docked the two Müller-Lyer figures aligned and thus created a combined stimulus with three sets of the wings (Fig. II.11-1d). In subsequent studies, the luminance or color contrast of the whole figures or their parts have been varied; gaps between shafts and wings were introduced; instead of the wings, the appendages of different shapes were used (Fig. II.11-1b, II.11-e, II.11-f, II.11-g); or the wings were defined by the ends of additional parallel lines (Fig. II.11-1c). The figure parts have been presented together or separately to different eyes, either simultaneously or alternately; stimuli has been formed by textures and became apparent only in binocular fusion; and the cumulative influence on the illusion magnitude of combinations of different illusory figures have been investigated.

It is evident that the illusions of the Müller-Lyer type are directly related to the performance of such routine tasks as the visual assessment of the linear extent of objects or their mutual localization. Therefore, the study of illusions can substantially narrow the range of theoretical suggestions regarding the relevant neural organization and contribute to a better understanding of principles underlying the perception of spatial relationships. Unfortunately, despite the enormous body of accumulated experimental data, the number of theoretical interpretations has not decreased and there remains no generally accepted explanation of the phenomenon.

In this connection, it remains an open question whether the same visual mechanisms are responsible for misperception of different modifications of stimuli. It is possible that for figures consisting of separate elements, the neural mechanisms that encode the coordinates of the elements are used, and the illusion is a result of errors in the coding. For figures comprising some combine components (e.g., shaft-lines) the mechanisms directly responding to the linear extent of these components can be involved, and, because of the nonmandatory transitivity of the geometric properties of the visual space (Gillam, 1998), each of the considered mechanisms can generate its own errors. Moreover, the illusions may be caused by errors in the simultaneous work of several mechanisms, operating at different levels of the visual system (Coren & Girgus, 1978). Thus, for various configurations of stimuli, the manifestation of the illusions can vary considerably and in an unpredictable way. Therefore, the search for a common set of principles to account for the widest possible range of effects arising in perception of stimuli of various modifications still remains quite topical. It is also important that these principles and the relationships between them be determined quantitatively, in order to clearly identify the limits of their use.

UNDERLYING PRINCIPLES BEHIND THE CENTROID-BASED APPROACH IN ILLUSIONS STUDY

Numerous searches of the relevant parameters determining the properties of the Müller-Lyer and related illusions have revealed one common regularity: illusions arise in the presence of contextual elements (distracters) located in the vicinity of stimulus terminators (shaft-ends or

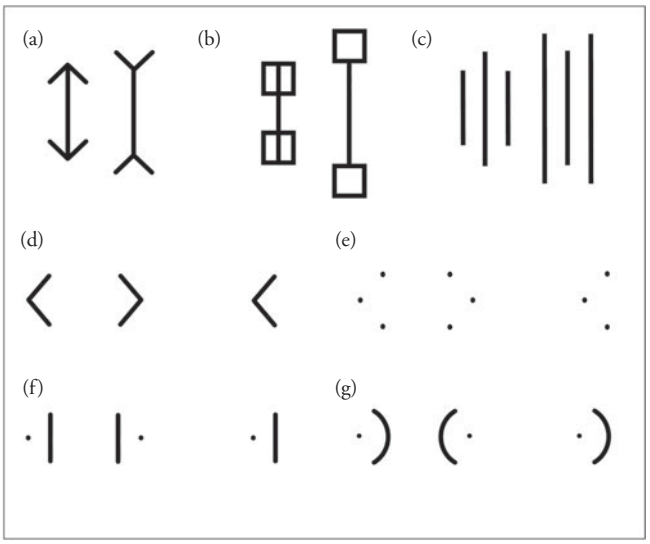


Figure II.11-1. Different types of stimuli that give rise to the illusion of extent. (a) Conventional illusory figures with the Müller-Lyer wings. (b) The wings are replaced by squares. (c) Imaginary wings are designated by the endpoints of parallel lines. Four modifications of the Brentano type of illusory figures comprising different contextual flanks: (d) the Müller-Lyer wings; (e) pairs of dots; (f) the vertical stripes; (g) arcs of a circle. In the latter three cases, the imaginary wings were formed by dots of stimulus terminators and dots or endpoints of adjacent flanks. (Author's own creation.)

delimiters of empty spatial intervals). It was found that even for a simplified figure composed of separate dots (Fig. II.11-1e), one can obtain quite stable and measurable effects. Such elementary stimulus has the typical features of the illusory figure and, at the same time, its application substantially reduces the number of parameters used in experiments (e.g., the mutual coordinates and relative brightness of the dots). This circumstance considerably facilitates the theoretical interpretation of the results obtained and offers another way to study illusions. Varying the stimulus parameters in a certain way—for example, lengthening or shortening the imaginary wings (Fig. II.11-2a), changing their internal angle (Fig. II.11-2b), altering the luminance of distracting dots (Fig. II.11-2c), or varying the length of shaft interval (Fig. II.11-2d)—makes it possible to establish their influence on illusion magnitude and helps us develop an appropriate quantitative description of the phenomenon. By introducing additional stimuli elements such as segments of lines (Fig. II.11-1d, II.11-f) or curves (Fig. II.11-1g), one can check whether the illusion effects possess additivity and to what extent the superposition principle can be applied. Gradually complicating the stimuli, one can try to extend the interpretation of regularities established for elementary shapes onto an increasingly wide range of images. Such a step-by-step development of a quantitative description of illusion can

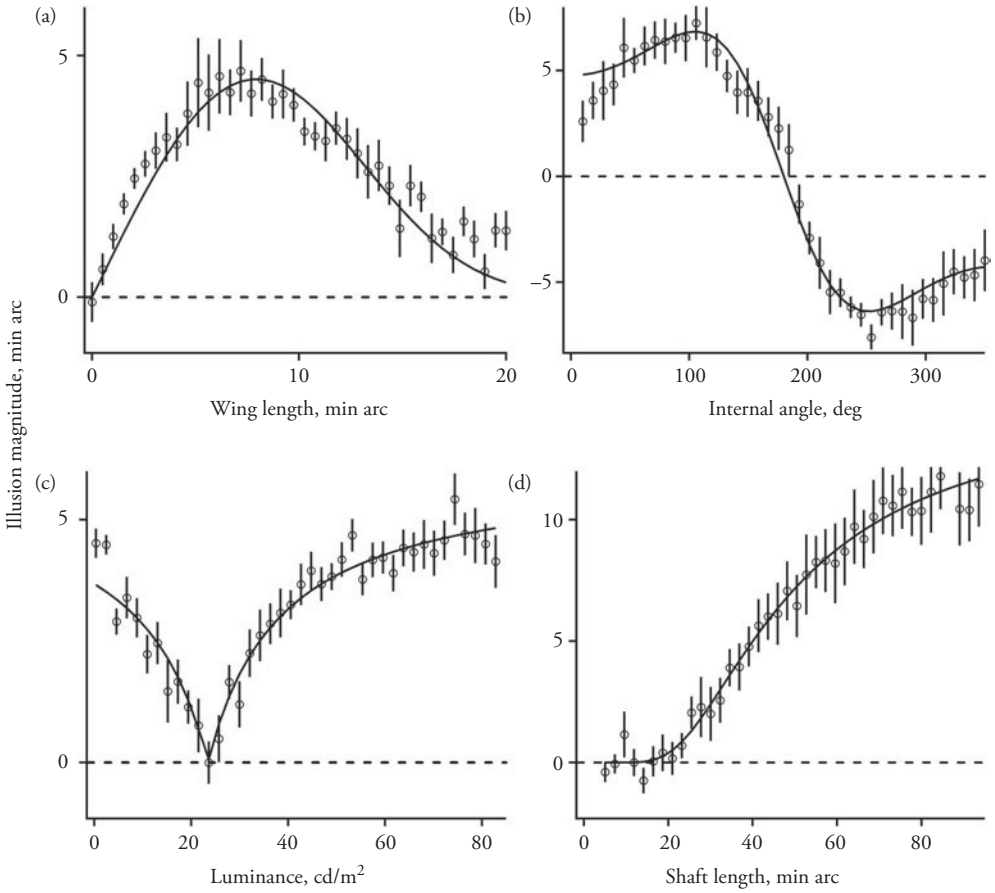


Figure II.11-2. Dependencies of the illusion strength on different parameters of the stimulus shown in Figure I.11-1e (from Bulatov, Bertulis, Bulatova, et al., 2009; Bulatov, Bertulis, Bielevičius, & Bulatova, 2009; Bulatov, Bertulis, Gutauskas, & Bulatova, 2010). Illusion magnitude as a function of the length (a) and internal angle (b) of imaginary wings. (c) Illusion magnitude as a function of the luminance of distracting dots. (d) Illusion magnitude as a function of the length of the shaft interval. Circles: The experimental data (*error bars*: ± one standard error of the mean; *solid curves*: the least squares fitting by model functions). (Adapted with permission from Bulatov, Bertulis, Bulatova, et al., 2009; Bulatov, Bertulis, Bielevičius, & Bulatova, 2009; Bulatov, Bertulis, Gutauskas, & Bulatova, 2010.)

provide immediate experimental verification of its predictions: besides the qualitative trends, the functions of an adequate model must also offer the quantitative assessments of illusion behavior for different variations of stimuli parameters.

As an example of such a quantitative approach in illusions study, the computational model based on the so-called centroid hypothesis can be considered. According to the hypothesis (Badcock, Hess, & Dobbins, 1996; McGraw, Whitaker, Badcock, & Skillen, 2003; Morgan & Glennerster, 1991; Seizova-Cajic & Gillam, 2006), judgments of distance between visual objects are strongly affected by the neural computation of centers-of-masses (centroids) of the luminance profiles of the objects (i.e., the centroid of an image determines its perceived location). As suggested (Morgan, Hole, & Glennerster, 1990), the procedure of centroid extraction is causally related to a certain integration of neural activations (spatial pooling of the positional signals) evoked by the neighboring object parts. Though the integration reduces the ability of the visual system to determine accurately the mutual position of image elements (i.e., coarsens the positional acuity), such pooling seems to be quite necessary and biologically substantiated since it allows fast and reliable assessment of the location of the visual object as whole, irrespective of its size, the shape complexity, and illumination conditions (coarse positional resolution is assumed as the particular feature of centroid extraction and therefore does not affect the overall visual acuity, and the perceived objects are not blurred). Concerning the Müller-Lyer and similar illusions, the pattern of neural excitation evoked by contextual flank (e.g., the wings themselves) overlaps with that caused by the stimulus terminator (e.g., the wings apex), thereby leading (due to the shift of the centroid of summed excitation) to its perceptual displacement. The relative displacement of all stimulus terminators leads to misjudgment of distances between them; that is, the illusion occurs as a side effect due to necessarily low spatial resolution of the neural mechanism of assessment of the relative location of the visual objects.

Thus the centroid-based explanation supposes an essentially local mechanism of the illusion occurrence, and it has been confirmed experimentally (Morgan et al., 1990; Post, Welch, & Caufield, 1998; Welch, Post, Lum, & Prinzmetal, 2004) that the effects of contraction/expansion are not distributed evenly along the shafts of the Müller-Lyer figures but are concentrated in areas immediately surrounding the wings–shaft intersections. The results of experiments with addition of nontarget dots to the conventional Müller-Lyer figures (Searleman, Porac, Dafoe, & Hetzel, 2005) or to the modified Brentano figure consisting of separate dots (Bulatov, Bertulis, Gutauskas, Mickiene, & Kadziene, 2010) also support the suggestion.

Since the centroid hypothesis implies the perceptual shifts of the stimulus terminators toward the centers-of-masses of distracters, for the figures with tilted flanks (i.e., their centroids do not coincide with the stimulus axis) the illusion magnitude should be proportional to the cosine function of the angle of tilting (the projection of the terminator's actual positional shift onto the axis). The rotation of flanks of any shape around the corresponding terminators should evoke the cosine modulation of illusion magnitude, and such a pattern of changes was confirmed in experiments with modified Brentano figures comprising the Müller-Lyer wings or arcs of a circle (Bulatov, Bertulis, Mickiene, Surkys, & Bielevicius, 2011).

Taking into account these considerations, the quantitative model of automatic centroid extraction has been developed (Bulatov, Bertulis, Bulatova, & Loginovich, 2009; Bulatov, Bertulis, Gutauskas, Mickiene, et al., 2010; Bulatov, Mickiene, Bulatova, & Kadziene, 2011). The computational procedure of the model consists of two sequential stages: (a) a weighted spatial pooling of the neural excitations evoked by the stimulus terminator and adjacent contextual flank and (b) the calculation of the position of the pooling centroid.

The suggestions concerning the first stage are based on the results of electrophysiological studies that have demonstrated the attentional modulation of the responses of neurons in a wide variety of brain areas (Desimone & Duncan, 1995; Posner & Petersen, 1990; Robinson & Petersen, 1992), as well as on evidences from functional magnetic resonance imaging experiments that have revealed enhancement of neural activation within a circumscribed region of the cortical representations of the attended stimulus (Fink, Marshall, Weiss, Toni, & Zilles, 2002; Gandhi, Heeger, & Boynton, 1999; Kelly, Gomez-Ramirez, & Foxe, 2008). As we know from psychophysical studies, the selective attention can diverge from gaze direction thus allowing simultaneous tracking of up to seven different targets (Bettencourt & Somers, 2009; Horowitz et al., 2007); also, the size of the relevant windows of attention increases with their visual field eccentricity (Intriligator & Cavanagh, 2001; Müller, Mollenhauer, Rösler, & Kleinschmidt, 2005; Nakayama & Mackaben, 1989; Toet & Levi, 1992). Therefore, it was assumed in the model that the procedure of the positional pooling can be described as the multiplication of the relief of the neural excitation by a circular Gaussian or a Mexican hat–like profile of certain attentional windows (Fig. II.11-3a). It is also assumed that the windows (whose sizes depend on the actual direction of the observer's gaze) are centered at stimulus terminators.

The second stage in the modeling can be represented as the spatial filtering (formally described in terms of two-dimensional convolution) that occurs in the system of evenly distributed receptive fields of certain summation units. For a physically accurate assessment of the centroid position on some arbitrary axis of the pattern, it is necessary for the weighting function of the receptive field to be in the form of an upside-down quadratic parabola along this axis and remain a constant in an orthogonal direction (Fig. II.11-3b, left). In the visual system, the neurons with the elongated receptive fields (Fig. II.11-3b, center), built up from aligned in a row subfields, each of which possesses the profile of a circular Gaussian, can be considered as candidates for the role of the summation units. Convolution of the weighting function of the receptive field with the profile of the pooled excitation yields the position of its centroid (Fig. II.11-3b, right) corresponding to the position of the most excited summation unit (i.e., to the locus of the peak of the envelope of the filtered excitation). In accordance with the centroid hypothesis, the visual system identifies the position of the pooling centroid with

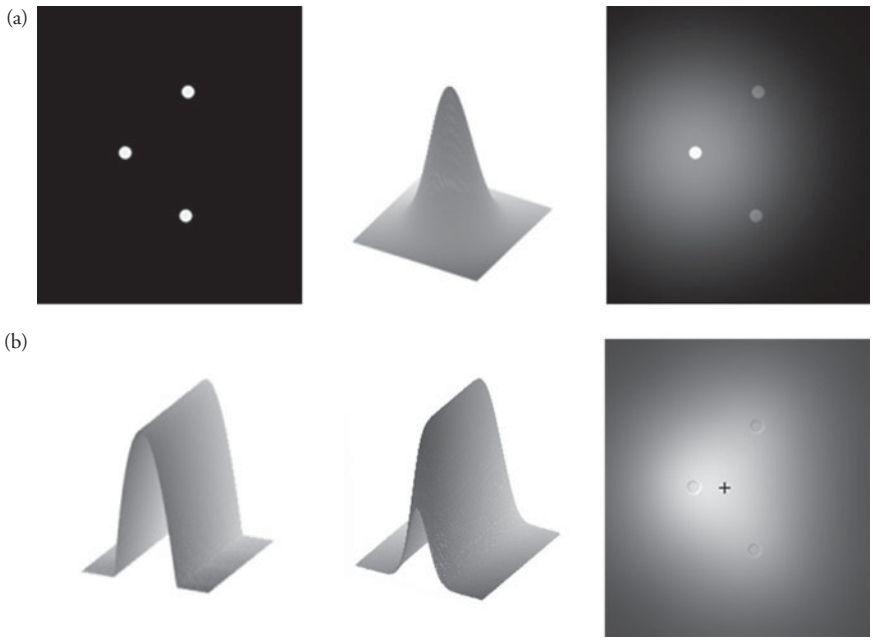


Figure II.11-3. Diagrams illustrating the modeling. (a) *Left*, a single cluster of dots of the stimulus shown in Fig. I.11-1e; *center*, circular Gaussian profile of the attentional pooling window; *right*, the result of applying of the pooling procedure (for illustrative purposes, the Gaussian distribution of the grey-level intensity indicates a two-dimensional envelope of attentional window). (b) Two different types of the weighting function of the receptive field of the summation units: *left*, the inverted quadratic parabola and *center*, the Gaussian; *right*, the resulting pattern of excitation obtained after applying of the convolution procedure (*cross* and *grey blobs*: the location of the pooling centroid and elements of the original stimulus pattern, respectively). (Author's own creation.)

that of the stimulus terminator, and the latter appears to be perceptually shifted.

From the coincidence of the calculated and experimental curves (Fig. II.11-2), one can conclude that the presented model rather well predicts the magnitude of the illusion of extent for the simple figure formed of separate dots. The use of figures with contextual flanks consisting of vertical stripes, arcs of a circle, or the Müller-Lyer wings also yields good agreement between the calculations and experimental data (Fig. II.11-4). Thus, at least for figures composed of separate clusters of elements, the principles of the modeling based on the suggestion that the illusions are caused by the losses in positional acuity (due to coding of stimulus parts locations via centroids) can be considered as quite substantiated. The task of further modeling studies is to test the applicability of these principles to account for an increasingly wide range of stimuli, including those incorporating combined elements.

CENTROID-BASED EXPLANATION AND OTHER THEORIES OF GEOMETRIC ILLUSIONS

Since centroid modeling inherently reduces to some procedures of a spatial-frequency filtering, it can be attributed to so-called physiological type of theories; other examples are the models of low- or band-frequency filtering (Bulatov, Bertulis, & Mickiene, 1997; Bulatov & Bertulis, 2005; Di Maio & Lansky, 1998; Ginsburg, 1984; Sierra-Vázquez & Serrano-Pedraza, 2010), as well as explanations of illusions' occurrences by internal noises in the neural networks (Fermüller & Malm, 2004). The distinctive properties of the centroid model are that it includes the procedures of selective attention and clearly specifies the feature of the observable object (the center-of-masses of the corresponding neural excitation), which the visual system can use in determining the object's location. Followers of the physiological approach are inclined to believe that the illusions arise due to the peculiarities

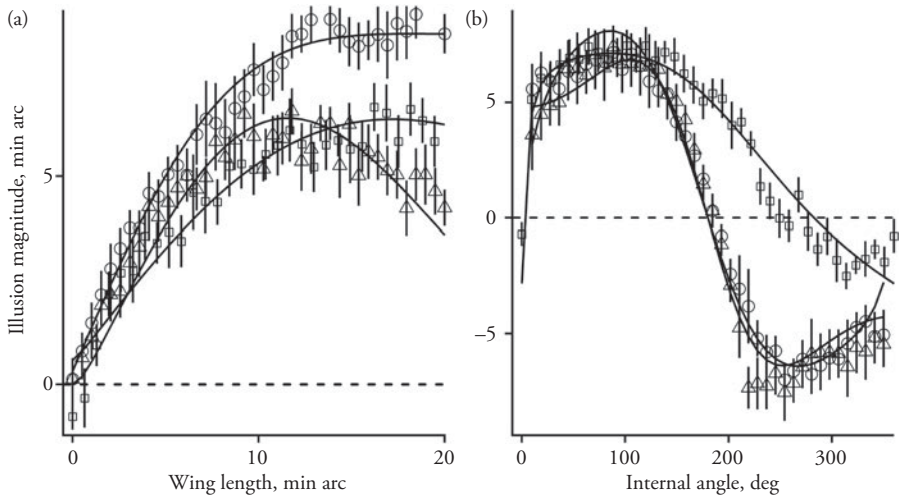


Figure II.11-4. Illusion magnitude as a function of the length (a) and internal angle (b) of the wings for stimulus shown in Figure I.11-1d, I.11-11f, and I.11-11g (from Bulatov, Bertulis, Bulatova, et al., 2009; Bulatov, Mickiene, et al., 2011). Contextual flanks consisted of the Müller-Lyer wings (*circles*), or vertical stripes (*triangles*), or arcs of a circle (*squares*). In latter two cases, the imaginary wings were formed by dots of stimulus terminators and endpoints of adjacent flanks. *Error bars*: ± one standard error of the mean; *solid curves*: the least squares fitting by model functions. (Adapted with permission from Bulatov, Bertulis, Bulatova, et al., 2009; Bulatov, Mickiene, et al., 2011.)

of neural processing of signals at relatively low levels of the visual system, and the contextual appendages of illusory figures appear as sources of interference that physically distort the signals evoked by the base elements of an image. It is considered to be enough for an explanation of illusion causes if, after implementation of a certain computational procedures, the profile of neural excitation becomes modified in such a way that reproduces the features of the investigated phenomenon (e.g., the distances between corresponding loci of the profile have changed properly). Hence, the existence of a certain higher-level neural mechanism that compares these distances is assumed implicitly; however, it remains unclear as to where in the brain this mechanism takes place and what the principles of its functioning are. Thus one of the common drawbacks of the physiological theories is disregard of the role of higher-level visual mechanisms in illusions emergence.

On the contrary, the adherents of the psychological approach are inclined to neglect the issues concerning low-level neural processing and consider the contextual constituents of the illusory figures as the elements of specific visual scenes, the subconscious interpretation of which leads to misjudgments of length. One of the most popular is the explanation based on assumption of misinterpretation of the flat drawings as to possess a certain three-dimensional structure (Gregory, 1968; Gillam, 1998; Redding & Parkinson, 2002; Redding & Vinson, 2010). It is supposed that the visual system is adapted for the proper perception of three-dimensional scenes only; therefore the Müller-Lyer wings appear for the observer as the sides of spatial angles, which causes the sensation of depth and leads to the perception of shaft-lines of the figures as more distant or closer; in turn, misapplication of the visual mechanism of size constancy gives rise to the illusion. According to other well-known psychological theories (e.g., Day, 2006), illusions may occur due to a tendency for the size of image parts to be affected by the size of the whole: the figure with the inward-pointing wings occupies less space than that with outward-pointing wings; therefore, if one figure is physically smaller, the size of its component (e.g., the shaft) also should be perceived as smaller. Similar to this explanation is the suggestion that the illusion strength depends on the size of the area of the empty space around the illusory figures (Nemati, 2009). According to authors of other popular conceptions (Howe & Purves, 2005), which to some extent can be also attributed to the psychological niche, the illusion can be explained by the statistical distributions of the geometric relationships in the environment and the knowledge of the probabilities is somehow preserved in perception. Unfortunately, psychological theories, despite their surprising ingenuity, typically give only qualitative predictions and, as a rule, do not report any link between the declared mechanisms of illusion occurrence and certain brain structures.

The fact that a generally accepted view concerning the illusions origin is still absent suggests that a new common approach is needed whereby physiological and psychological explanations supplement each other. It is highly probable that each of the currently existing theories contains its own part of truth, and further research will allow us to see how these pieces fit into a coherent and comprehensive explanation of the geometric illusions of extent.

REFERENCES

Badcock, D. R., Hess, R. F., & Dobbins, K. (1996). Localization of element clusters: Multiple cues. *Vision Research*, *36*, 1467–1472.

Bettencourt, K. C., & Somers, D. C. (2009). Effects of target enhancement and distractor suppression on multiple object tracking capacity. *Journal of Vision*, *7*, 1–11.

Brentano, F. (1892). Über ein optisches Paradoxen. *Zeitschrift für Psychologie*, *3*, 349–358.

Bulatov, A., & Bertulis, A. (2005). Distracting effect in length matching. *Acta Neurobiologiae Experimentalis*, *65*, 265–269.

Bulatov, A., Bertulis, A., Bielevičius, A., & Bulatova, N. (2009). Interpretation of illusions of extent based on the centroid conception. *Sensory Systems*, *23*(1), 3–12.

Bulatov, A., Bertulis, A., Bulatova, N., & Loginovich, Y. (2009). Centroid extraction and illusions of extent with different contextual flanks. *Acta Neurobiologiae Experimentalis*, *69*, 504–525.

Bulatov, A., Bertulis, A., Gutauskas, A., & Bulatova, N. (2010). Stimulus size and visual illusion of extent. *Human Physiology*, *36*(2), 50–58.

Bulatov, A., Bertulis, A., Gutauskas, A., Mickiene, L., & Kadziene, G. (2010). Center-of-mass alterations and visual illusions of extent. *Biological Cybernetics*, *102*, 475–487.

Bulatov, A., Bertulis, A., & Mickiene, L. (1997). Geometrical illusions: Study and modelling. *Biological Cybernetics*, *77*, 395–406.

Bulatov, A., Bertulis, A., Mickiene, L., Surkys, T., & Bielevicius, A. (2011). Contextual flanks' tilting and magnitude of illusion of extent. *Vision Research*, *51*, 58–64.

Bulatov, A., Mickiene, L., Bulatova, N., & Kadziene, G. (2011). Center-of-mass' shift and the curvilinear distractors in visual illusion of extent. *Sensory Systems*, *25*(1), 45–54.

Coren, S., & Girgus J. S. (1978). *Seeing is deceiving: The psychology of visual illusions*. Hillsdale, NJ: Lawrence Erlbaum.

Day, R. H. (2006). Two principles of perception revealed by geometrical illusions. *Australian Journal of Psychology*, *58*(3), 123–129.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.

Di Maio, V., & Lansky, P. (1998). The Müller-Lyer illusion in interpolated figures. *Perceptual and Motor Skills*, *87*, 499–504.

Fermüller, C., & Malm, H. (2004). Uncertainty in visual processes predicts geometrical optical illusions. *Vision Research*, *44*, 727–749.

Fink, G. R., Marshall, J. C., Weiss, P. H., Toni, I., & Zilles, K. (2002). Task instructions influence the cognitive strategies involved in line bisection judgments: Evidence from modulated neural mechanisms revealed by fMRI. *Neuropsychologia*, *40*, 119–130.

Gandhi, S. P., Heeger, D. J., & Boynton, G. M. (1999). Spatial attention affects brain activity in human primary visual cortex. *Proceedings of the National Academy of Sciences USA*, *16*, 3314–3319.

Gillam, B. (1998). Illusions at century's end. In J. Hochberg (Ed.), *Perception and cognition at century's end* (pp. 95–136). San Diego, CA: Academic Press.

Ginsburg, A. P. (1984). Visual form perception based on biological filtering. In L. Spilman & G. R. Wooten (Eds.), *Sensory experience, adaptation and perception* (pp. 53–72). Hillsdale, NJ: Lawrence Erlbaum.

Gregory, R. L. (1968). Visual illusions. *Scientific American*, *219*, 66–67.

Horowitz, T. S., Klieger, S. B., Fencsik, D. E., Yang, K. K., Alvarez, G. A., & Wolfe, J. M. (2007). Tracking unique objects. *Perception & Psychophysics*, *69*, 172–184.

Howe, C. Q., & Purves, D. (2005). The Müller-Lyer illusion explained by the statistics of image-source relationships. *Proceedings of the National Academy of Sciences USA*, *102*, 1234–1239.

Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology*, *43*, 171–216.

Kelly, S. P., Gomez-Ramirez, M., & Foxe, J. J. (2008). Spatial attention modulates initial afferent activity in human primary visual cortex. *Cerebral Cortex*, *18*, 2629–2636.

McGraw, P. V., Whitaker, D., Badcock, D. R., & Skillen, J. (2003). Neither here nor there: Localizing conflicting visual attributes. *Journal of Vision*, *3*, 265–273.

Morgan, M. J., & Glennerster, A. (1991). Efficiency of locating centers of dot-clusters by human observers. *Vision Research*, *31*, 2075–2083.

Morgan, M. J., Hole, G. J., & Glennerster, A. (1990). Biases and sensitivities in geometrical illusions. *Vision Research*, *30*, 1793–1810.

Müller, N. G., Mollenhauer, M., Rösler, A., & Kleinschmidt, A. (2005). The attentional field has a Mexican hat distribution. *Vision Research*, *45*, 1129–1137.

Müller-Lyer, F. C. (1889). Optische Urtheilstäuschungen. *Archiv für Anatomie und Physiologie*, *2*(Suppl.), 263–270.

Nakayama, K., & Mackaben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research*, *29*, 1631–1647.

Nemati, F. (2009). Size and direction of distortion in geometric-optical illusions: Conciliation between the Müller-Lyer and Titchener configurations. *Perception*, *38*(11), 1585–1600.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*, 25–42.

Post, R. B., Welch, R. B., & Caufield, K. (1998). Relative spatial expansion and contraction within the Müller-Lyer and Judd illusions. *Perception*, *27*, 827–838.

Redding, G. M., & Parkinson, S. L. (2002). Contributions of virtual corner size and picture plane size to the Müller-Lyer fork junction illusion. Paper presented at the meeting of the Midwestern Psychological Association, Chicago, May.

Redding, G. M., & Vinson, D. W. (2010). Virtual and drawing structures for the Müller-Lyer illusions. *Attention, Perception & Psychophysics*, *72*(5), 1350–1366.

Robinson, D. L., & Petersen, S. E. (1992). The pulvinar and visual salience. *Trends in Neurosciences*, *15*, 127–132.

Searleman, A., Porac, C., Dafoe, C., & Hetzel, B. (2005). Altering Mueller-Lyer illusion magnitude using figural additions at the wing–shaft intersections. *American Journal of Psychology*, *118*, 619–637.

Seizova-Cajic, T., & Gillam, B. (2006). Biases in judgments of separation and orientation elements belonging to different clusters. *Vision Research*, *46*, 2525–2534.

Sierra-Vázquez, V., & Serrano-Pedraza, I. (2010). Application of Riesz transforms to the isotropic AM-PM decomposition of geometrical-optical illusion images. *Journal of the Optical Society of America*, *27*(4), 781–796.

Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research*, *32*, 1349–1357.

Welch, R. B., Post, R. B., Lum, W., & Prinzmetal, W. (2004). The relationship between perceived length and egocentric location in Müller-Lyer figures with one versus two chevrons. *Perception & Psychophysics*, *66*, 1095–1104.

Chapter 12
The Bar-Cross-Ellipse Illusion

Gideon P. Caplovitz, Alex Boswell, and Kyle Killebrew

INTRODUCTION

As an object moves through the environment, it will frequently move behind other objects, thereby becoming partially occluded from view. In constructing the percept of such a partially occluded object in motion, the visual system must solve a distinct set of problems. Because the shape and global motion of the object is not fully represented in the visual scene, the brain must make assumptions about what the object may look like and how it may be moving. The task is made more difficult by the fact that there are two fundamental ambiguities that arise at the borders of occlusion: (a) the contour formed at the point of occlusion may belong to either the occluding or occluded object and (b) motion signals that arise at the point of occlusion may be due to the motion of either the occluded or occluding surface. Here we describe the bar-cross-ellipse illusion (Caplovitz & Tse, 2006) that reveals how the manner in which these ambiguities are resolved can dramatically alter our perceptual experience of an otherwise seemingly simple stimulus.

Illustrated in Figure II.12-1, the basic stimulus is constructed by first placing an ellipse on a uniform background and then placing four squares on top of the ellipse, leaving only a cross-shaped portion of the surface visible. Importantly, the squares must be the same color as the background; however, the ellipse can be any color that differs from the background and squares. If the ellipse is then made to continuously rotate (see Video II.12-1 on the Companion Website), prolonged viewing of the stimulus will lead an observer to perceive four distinct motion percepts, only one of which directly corresponds with how the stimulus is constructed. Over time, the four percepts will be perceived in alternating fashion. Each percept corresponds to a specific figure–ground segmentation formed by the assignment of contours and motion signals to specific objects in either the foreground or background. In the following sections we describe each percept in turn.

THE MORPHING CROSS

When naïve observers fist see the bar-cross-ellipse, the initial percept is often that of a single figure in the form of a cross, stretching and shrinking in the two-dimensional (2D) plane (Fig. II.12-2a). Here the square occluders are not perceived as such but are instead assimilated into the background. As such, all contours present in the image are assigned to the single cross-shaped figure. All motion signals in the image are also assigned to the figure and constrained to the 2D image plane.

BARS OSCILLATING IN DEPTH

Over time, this stimulus can be perceived as two distinct perpendicular bars oscillating in depth (Fig. II.12-2b). Again, the squares are not perceived as occluding surfaces, and all contours and motion signals are attributed to the figures formed by the exposed portions of the elliptical surface. However, here the inward and outward motions of the bar ends are interpreted as arising from three-dimensional (3D) changes in depth rather than a change in shape, as is the case in the morphing cross. Together, these two percepts reveal an intriguing relationship between 2D and 3D motions. Rigid motions in 3D will commonly generate nonrigid motions when projected into two dimensions (Wallach, 1953). There is evidence to suggest that, whenever possible, the visual system will construct a 3D rigid motion out of nonrigid 2D motion signals (Ullman, 1984). By segmenting the image into two distinct figures—that at times seem to paradoxically pass through each other in depth—this percept reveals the lengths the visual system will go to construct a rigid motion percept.

THE OCCLUDED ELLIPSE

Over time, the percept will eventually switch to that of an occluded ellipse rotating behind four square occluders. This is the percept that most directly corresponds to how the stimulus is constructed (Fig. II.12-1). Through the processes of amodal completion (Day & Kasperczyk, 1983) and anorthoscopic perception (Zöllner, 1862), the visual system constructs the global form and motion of a rotating ellipse from sparsely represented local form and motion information. In addition to integrating the small, exposed portions of contour to create the form of an ellipse, this percept is based on the differential assignment of two distinct sources

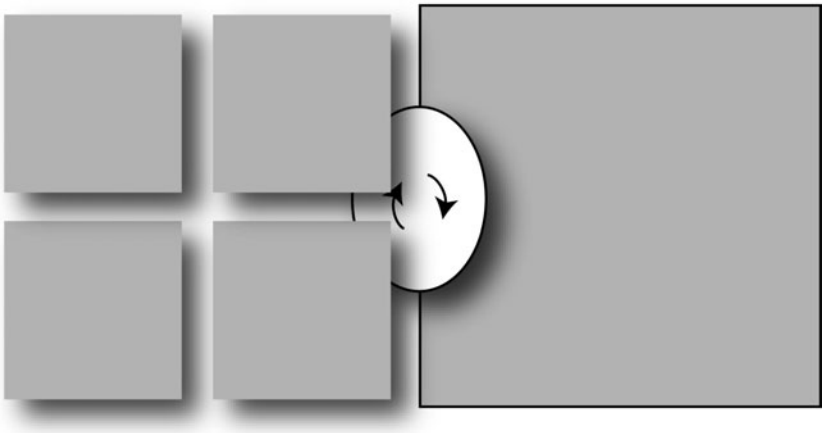
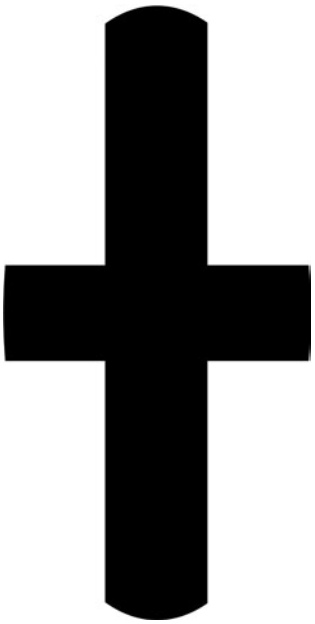


Figure II.12-1. How the bar-cross-ellipse stimulus is constructed. (Gideon Caplovitz)



Video II.12-1. The Bar-Cross-Ellipse. (Gideon Caplovitz)

of motion information: (a) the component motion (Adelson & Movshon, 1982), which is orthogonal to the small portions of the exposed elliptical contour, and (b) the feature motion, which is parallel to the contours formed at the occluding edges of the squares (Nakayama & Silverman, 1988a,b). In this percept, the component motions arising along the exposed portions of the ellipse are attributed to the ellipse itself. Presumably, these motion signals are integrated across the spatially separated portions of contour to support the percept of rotational motion. In contrast, the feature motions that arise along the occlusion contours are seemingly discarded, having been determined as arising due to occlusion (Nakayama & Silverman, 1988a,b), and do not contribute to the overall motion percept. Looking back at the previous percepts, it is now clear that the feature motions that are discarded here play a significant role in the 2D motion perceived in the morphing cross and the 3D motions in the bars oscillating in depth.

THE ROTATING ELLIPTICAL APERTURE

This is perhaps the most counterintuitive and perceptually striking of the percepts. Over time, the stimulus can be perceived such that a stationary cross is viewed through an elliptical aperture that is itself rotating (Fig. II.12-2c). Here, the contours formed by the occluding squares are assigned to a single cross-shaped figure, whereas the contours of the exposed portions of the ellipse are assigned to the foreground aperture. None of the motion signals present in the image are attributed to the cross figure. As in the occluded ellipse percept, the feature motions are again deemed to be spurious and are discarded. However, these feature-motion signals are attributed to the occlusion of the cross by a rotating foreground elliptical aperture. In the case of the occluded ellipse, the feature-motion signals are attributed to the occlusion of a rotating ellipse by stationary foreground squares. The rotational motion of the elliptical aperture is constructed by assigning the component motion signal present at the ends of the exposed contours to a foreground elliptical aperture. What is so fascinating about this percept is that the visual system constructs an illusory contour (Day & Kasperczyk, 1983) of the elliptical aperture as it rotates. When observing this percept, the illusory contour can be seen by examining the space covered by any one of the squares.

VARIANTS OF THE BAR-CROSS-ELLIPSE ILLUSION

As already stated, in general, when naïve observers view Video II.12-1, the initial percept is that of a morphing cross. Over time, the other percepts are seen and will typically switch back and forth between each other, as is the case with other well-known, multistable percepts (Necker, 1832; Rubin, 1921). It is our experience that once all four percepts have been observed, they tend to switch more frequently and the initial percept of the morphing cross becomes less common. Indeed, having viewed the stimulus so much, it is difficult for the first author to experience the morphing cross at all. It appears that having established that more complex, multilayered figure–ground segmentations can exist, the visual system is biased toward producing them. It is also worth noting that of the four percepts,

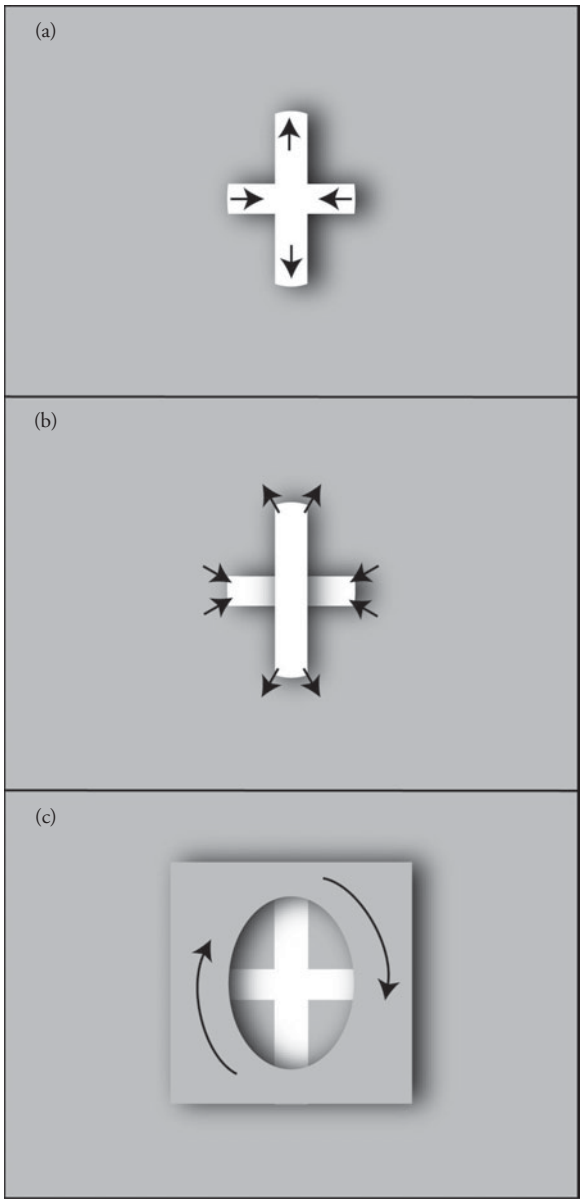
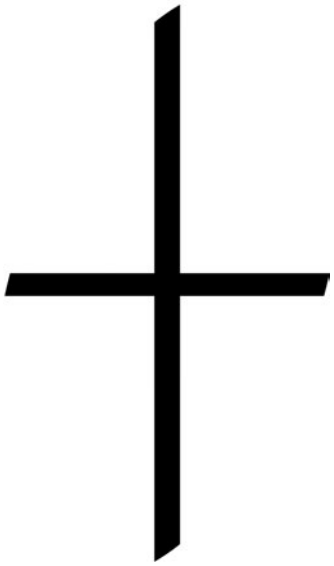


Figure II.12-2. Alternate percepts of the bar-cross-ellipse, the morphing cross, bars oscillating in depth, and a rotating elliptical aperture. (Gideon Caplovitz)

only the morphing cross corresponds to a nonrigid motion, and it may be that given the option of constructing a rigid motion the visual system will do so (Ullman, 1984).

Making small changes to the stimulus can greatly influence whether or how frequently each percept is perceived. For example, increasing the size of the occluding squares so that much less of the elliptical surface is visible (see Video II.12-2 on the Companion Website) makes it more difficult to experience either of the two elliptical percepts. However, with a trained eye, both of these percepts can still be perceived.

Intriguingly, making the ellipse skinnier (see Video II.12-3 on the Companion Website), facilitates seeing the occluded ellipse percept, whereas making the ellipse fatter makes it more difficult to see either of the ellipse percepts (see Video II.12-4 on the Companion Website). This may be



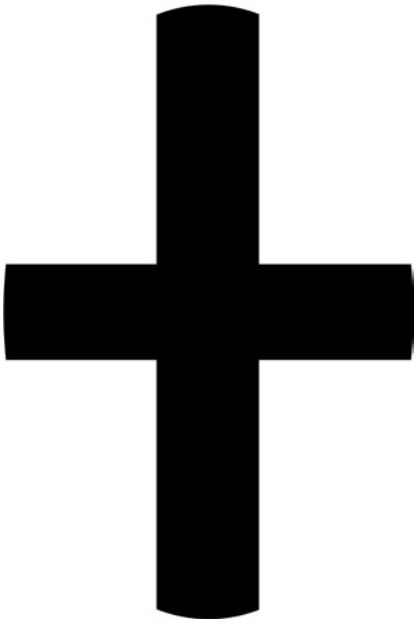
Video II.12-2. The Bar-Cross-Ellipse: Narrow Apertures. (Gideon Caplovitz)

the case because when the regions of high curvature at the ends of the skinny ellipse are exposed they provide stronger form cues for integrating the elliptical contour (Attneave, 1954).

The bar-cross-ellipse stimulus is quite similar in design to that underlying the breathing square Illusion (see Video II.12-5 on the Companion Website), in which a rotating square is used instead of a rotating ellipse (Shiffrar & Pavel, 1991). Despite their similarities, the percept of the breathing square stimulus is quite distinct from that of the bar-cross-ellipse. For one, the dominant percept is that of an occluded rotating square (analogous to the occluded ellipse percept); however, the square appears to loom in and out of depth as it rotates. It appears that the feature motions that are discarded in the



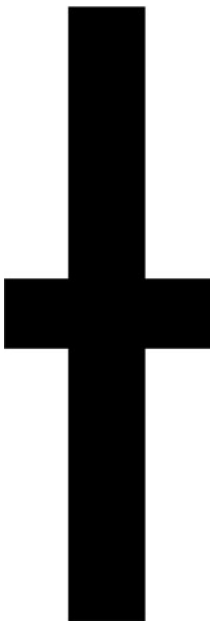
Video II.12-3. The Bar-Cross-Ellipse: Skinny Ellipse. (Gideon Caplovitz)



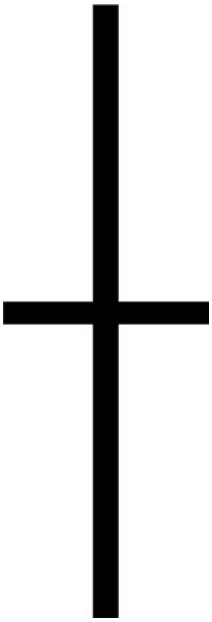
Video II.12-4. The Bar-Cross- Ellipse: Fat Apertures. (Gideon Caplovitz)



Video II.12-5. The Breathing Square. (Gideon Caplovitz)



Video II.12-6. The Bar-Cross-Rectangle. (Gideon Caplovitz)



Video II.12-7. The Bar-Cross-Rectangle: Narrow Apertures. (Gideon Caplovitz)

occluded ellipse percept are not discarded in the breathing square and are instead attributed to the contours of the figure and correspond to rigid motions in depth. As such, the precept of the breathing square appears to be an amalgamation of both the occluded ellipse and bars oscillating in depth percepts. We note that the percept of the breathing square is quite stable, and it is difficult if not impossible to perceive the stimulus as a morphing cross or rotating square aperture.

Here we describe for the first time a novel variant of the bar-cross-ellipse stimulus in which the ellipse is replaced by a rectangle (see Video II.12-6 on the Companion Website). Unlike the case of the breathing square, the percept here is multistable, and with prolonged viewing and a trained eye, percepts corresponding to each of the four bar-cross-ellipse percepts can be observed. Note that it is difficult to see this stimulus as a morphing cross or rotating aperture, perhaps because the corners of the rectangle provide such a strong cue for spatial integration (Attneave, 1954). The most striking difference is the biphasic nature of rotational speed when the percept is that of either a rotating occluded rectangle or a rotating rectangular aperture. As the rectangle rotates, it appears to speed up and slow down in a dramatic fashion. Increasing the aspect ratio of the rectangle (see Video II.12-7 on the Companion Website) accentuates this effect to the point where the rectangle almost appears to stop entirely.

NEURAL MECHANISMS UNDERLYING THE BAR-CROSS-ELLIPSE ILLUSION

Neural mechanisms underlying the bar-cross-ellipse illusion itself have not yet been directly investigated. However, a great deal of electrophysiological work, recording from neurons in nonhuman primates, has been performed on some of the illusion's important component processes, most notably the formation of illusory contours and the assignment of contour ownership.

Neurons that process illusory contours are found early in the visual processing stream; however, there is some debate as to just how early. In the most widely cited study, von der Heydt, Peterhans, and Baumgartner (1984) reported finding neurons in V2, but not V1, that responded to illusory contours. In contrast, Lee and Nguyen (2001) found neurons in V1 that did respond to illusory contours. Neurons have been found by Zhou, Friedman, and van der Heydt (2000) in V1, V2, and V4 that respond to oriented luminance edges, specifically edges that extend beyond their receptive fields. These responses are modulated by which side the figure is perceived to be on, and the neurons encode which side the edge belongs to. These neurons appear to belong to pairs of competing neural populations: one set of neurons are activated while the other is inhibited, depending on which of the two surfaces creating the edge belongs to the figure. Craft, Schutze, Niebur, and von der Heydt (2007) proposed a neural model in which these neurons (border-ownership neurons) project to a second kind of neuron, called grouping neurons. Grouping neurons send feedback to the border-ownership neurons, influencing their activity though excitation or inhibition. This interaction forms a basis for strengthening the signals of border ownership cells that are representing common figure–ground segmentations. Presumably these interactions proceed in a hierarchical fashion until a complete figure can be segmented from the background, thereby leading to a complete representation of the object. The multistable percepts of the bar-cross-ellipse stimulus seem to suggest that multiple populations of such border-ownership neurons, each associated with a given percept, may be simultaneously active and competing with each other for subjective representation. As we switch from one percept to the next, it is easy to imagine a distinct population of cells exceeding some sort of threshold and taking over the representational priority from another population (Kang & Blake, 2010). What we find remarkably intriguing in the potential interaction between border ownership and illusory contour cells revealed by switches between the occluded ellipse and rotating elliptical aperture percepts. Each of these percepts corresponds to a complimentary assignment of border ownership at the ends of the exposed elliptical contour. Does the switch of border ownership to the rotating elliptical aperture percept induce the illusory contours (i.e., release from inhibition the illusory contour cells)? Or does the formation of illusory contours induce the switch of border ownership?

While these and other questions remain unknown for the time being, we are reminded once again how simple stimuli can lead to visual illusions that highlight the constructive processes underlying visual perception in aesthetic ways that are easy to appreciate.

REFERENCES

Adelson, E. H., & Movshon J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, *300*, 523–525.

Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, *61*, 183–193.

Caplovitz, G. P, & Tse, P. U. (2006). The bar-cross-ellipse illusion: Alternating percepts of rigid and non-rigid motion based on contour ownership and trackable feature assignment. *Perception*, *35*, 993–997.

Craft, E., Schutze, H., Niebur, E., & von der Heydt, R. (2007). A neural model of figure–ground organization. *Journal of Neurophysiology*, *97*, 4310–4326.

Day, R., & Kasperczyk, R. (1983). Amodal completion as a basis for illusory contours. *Attention, Perception, & Psychophysics*, *33*, 355–364.

Kang, M., & Blake, R. (2010). What causes alternations in dominance during binocular rivalry? *Attention, Perception, & Psychophysics*, *72*(1), 179–186.

Lee, T. S., & Nguyen, M. (2001). Dynamics of subjective contour formation in the early visual cortex. *Proceedings of the National Academy of Science USA, 98*(4), 1907–1911.

Nakayama, K., & Silverman, G. H. (1988a). The aperture problem: I. Perception of nonrigidity and motion direction in translating sinusoidal lines. *Vision Research*, *28*(6), 739–746.

Nakayama, K., & Silverman, G. H. (1998b). The aperture problem: II. Spatial integration of velocity information along contours. *Vision Research*, *28*(6), 747–753.

Necker, L. A. (1832). Observations on some remarkable optical phaenomena seen in Switzerland; and on an optical phaenomenon which occurs on viewing a figure of a crystal or geometrical solid. *London and Edinburgh Philosophical Magazine and Journal of Science*, *5*, 329–337.

Rubin, E. (1921). *Visuell wahrgenommene figuren: Studien in psychologischer Analyse*. Kobenhaven: Gyldendalske boghandel.

Shiffrar, M., & Pavel, M. (1991). Percept of rigid motion within and across apertures. *Journal of Experimental Psychology: Human Perception & Performance*, *17*, 749–761.

Ullman, S. (1984). Maximizing rigidity: The incremental recovery of 3-D structure from rigid and nonrigid motion. *Perception*, *13*, 255–274.

von der Heydt, R., Peterhans, E., & Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science*, *224*(4654), 1260–1262.

Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, *45*, 205–217.

Zhou, H., Friedman, H. S., & von der Heydt, R. (2000). Coding of border ownership in monkey visual cortex. *Journal of Neuroscience*, *20*, 6594–6611.

Zöllner, F. (1862). Uë ber eine neue Art anorthoskopischer Zerrbilder. *Annalen der Physik und Chemie*, *117*, 477–484.

Chapter 13
The Spinning Ellipse Speed Illusion

Gideon Paul Caplovitz, Po-Jang Hsieh, Peter J. Kohler, and Katharine B. Porter

INTRODUCTION

How the perceived motion of an object is constructed from the spatiotemporal dynamics of the retinal image is a fundamental question that continues to challenge vision scientists. This is true even for the perception of relatively simple moving objects such as basic geometric shapes. At the heart of the problem is the fact that an infinite number of three-dimensional object motions can generate the same sequence as two-dimensional retinal images. The problem of interpreting this many-to-one mapping is commonly termed the 'inverse problem' or the 'aperture problem' in motion perception (Adelson & Movshon 1982; Fennema & Thompson 1979; Marr 1982; Nakayama & Silverman 1988). While the manner in which the aperture problem is solved remains unknown, the fact that it exists highly constrains how accurately an object's motion can be perceived (Stumpf, 1911/1996; Wallach, 1935).

The spinning ellipse speed illusion illustrates how the aperture problem, and the way in which the visual system solves it, influences the perceived speed of a rotating object. The spinning ellipse speed Illusion refers to the subjective experience that a low-aspect ratio 'fat' ellipse will appear to rotate more slowly than a higher-aspect ratio 'skinny' ellipse that is rotating at the same speed (Caplovitz, Hsieh, & Tse 2006; Caplovitz & Tse, 2007a; Kohler, Caplovitz, Hsieh, Sun, & Tse, 2010; Porter, Caplovitz, Kohler, Ackerman, & Tse, 2011). The illusion is illustrated in Figure II.13-1a and can be observed in Video II.13-1 on the Companion Website. The fundamental contribution of the illusion to the understanding of visual perception is that it demonstrates that the shape of an object influences the speed with which it appears to rotate. This basic observation challenges the classic belief that the form and motion of an object are processed independently. Specifically, objects defined by smooth, closed contours with low curvature appear to rotate more slowly than those whose contours have regions of high or discontinuous curvature (such as sharp corners). The following sections describe the conditions under which the illusion is observed and how the illusion relates to the aperture problem.

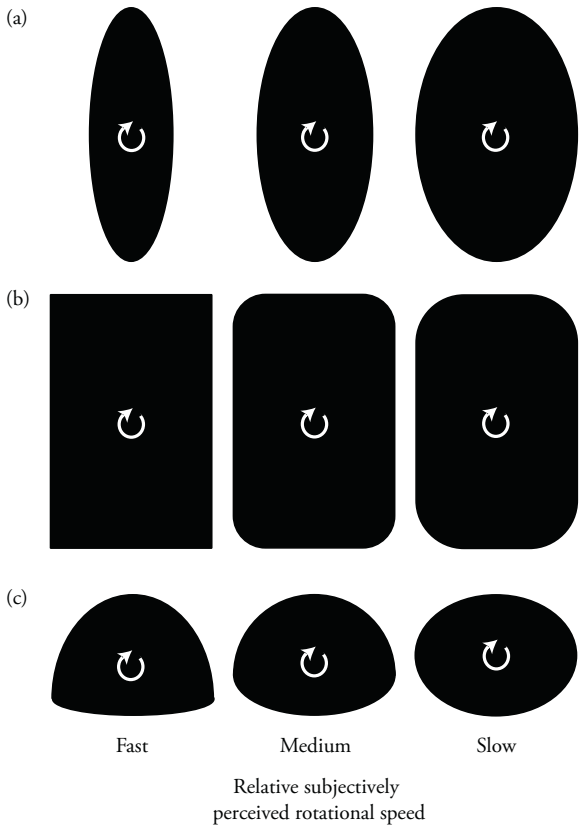


Figure II.13-1. Stimuli whose contours have regions of high or discontinuous curvature will appear to rotate faster than those that do not. This is true for (a) ellipses, (b) bumps, and (c) rounded rectangles.

THE SPINNING ELLIPSE SPEED ILLUSION MANIFESTS ACROSS A WIDE RANGE OF STIMULUS CONDITIONS

As illustrated in Video II.13-1 on the Companion Website, the spinning ellipse speed Illusion is a relatively strong effect that can be observed with free viewing in largely uncontrolled viewing conditions. Here we briefly review the stimulus conditions under which we have measured the illusion, including the manner in which the elliptical shapes



Video II.13-1. Here the skinny ellipse will appear to rotate faster than the fat one. In fact they are rotating at the same angular velocity.

are defined: their eccentricity, overall speed, and nonelliptical shapes.

In our initial investigation into the mechanisms that underlie the illusion (Caplovitz et al., 2006), we characterized the magnitude of the illusion for ellipses defined by luminance contrast relative to the background. We also found that the illusion can be observed and is just as strong if the ellipses are defined by chromatic contrast or motion contrast. Intriguingly, the contour of the ellipse need not be explicitly defined in order for the illusion to occur. We found that the illusion persists when using a set of small dots to form an elliptical "dotted" contour, as illustrated in Figure II.13-2. This was true even if the dots were contrast-balanced so as to be equiluminant with the background, indicating that low spatial frequency channels alone are not responsible for the presence of the speed illusion. However, the illusion will be perceived only for dotted contours if the dots are closely spaced together. If the dots are too widely spaced along the elliptical contour, aspect ratio will have no effect on the perceived speed of rotation (Caplovitz & Tse, 2007a). This suggests that the perceived angular velocity of a rotating object may arise from more than one source of motion information.

The influence of aspect ratio on the perceived speed of a rotating ellipse is largely invariant to both overall speed and eccentricity. We have measured the illusion with overall speeds ranging from very slow (4.2º/s; Kohler et al., 2010) to very fast (187º/s; Porter et al., 2011). In addition, we found the relative slowdown associated with making

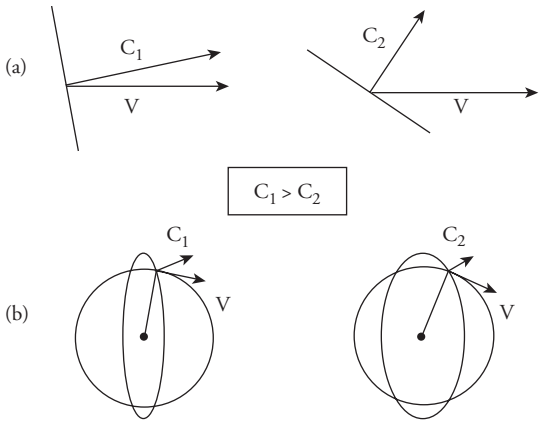


Figure II.13-2. Due to the receptive field properties of early motion detecting neurons, only the component of motion orthogonal to a moving contour is detectable. (a) The component motion C for a line translating with velocity V will be greater for the line on the left than the one on the right. (b) The component motion for a point along the contour of a rotating skinny ellipse (left) will be greater than that of a corresponding point along the contour of a rotating fat ellipse (right).

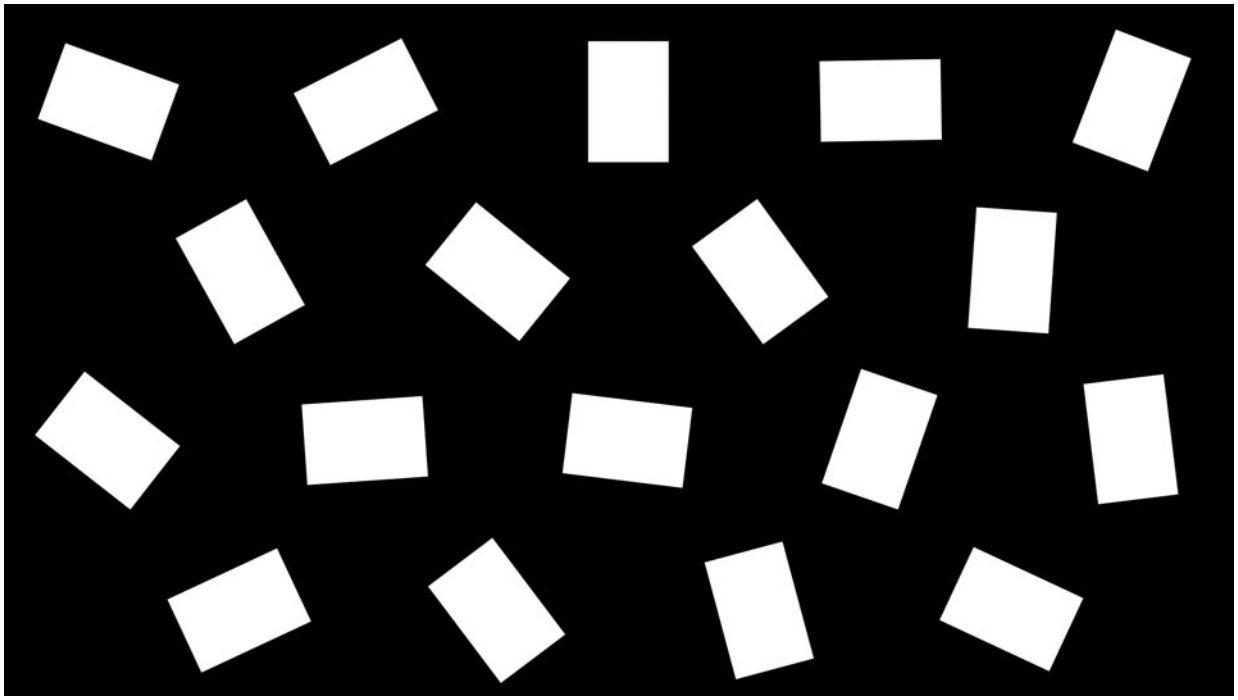
an ellipse fatter remains the same across eccentricities ranging from 6.9º to 16.3º visual angle (Porter et al., 2011). Furthermore, the illusion is observed even if the ellipses are simultaneously translating as they rotate about their centers, and the magnitude of the illusion is unaffected by the speed with which they are translating (Porter et al., 2011). Thus in the case of combined translational and rotational motion, the translational component of the motion appears to be processed independently from the rotational motion cues leading to the illusion.

Finally, the illusion is not limited to elliptical shapes. We have, for example, demonstrated that a rectangle will appear to rotate faster than another rectangle that has had its corners "rounded off" as illustrated in Figure II.13-1b and Video II.13-2 on the Companion Website. Replacing the corners with rounded regions of systematically lower curvature leads to the rectangle appearing to rotate even slower (Caplovitz et al., 2006). Objects formed by combining two half ellipses of differing aspect ratio ("bumps") as illustrated in Figure II.13-1c will exhibit an illusory slowdown. As can be observed in Video II.13-3 on the Companion Website, objects formed by two half ellipses with very different aspect ratios will appear to rotate faster than those with more similar aspect ratios (Caplovitz & Tse, 2007b).

THEORETICAL UNDERPINNINGS OF THE SPINNING ELLIPSE SPEED ILLUSION

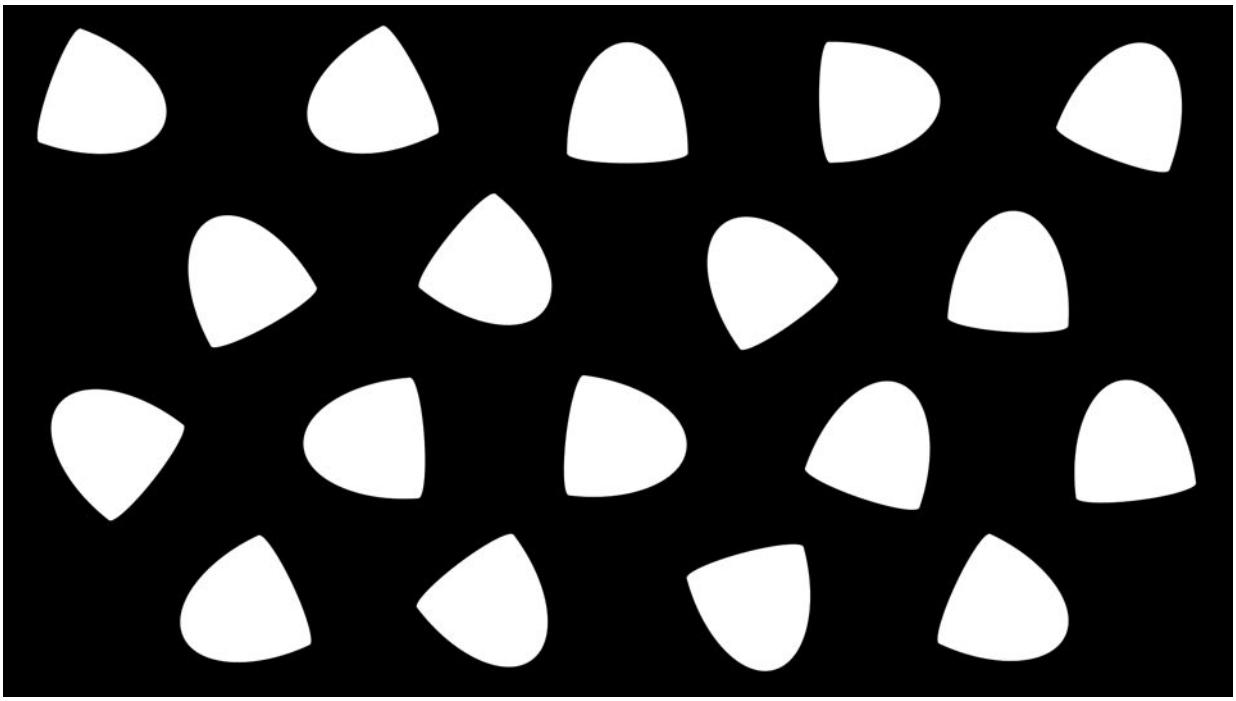
There are two nonmutually exclusive hypotheses for why the illusion exists, both related to the way in which the visual system may solve the aperture problem. The first has to do with how local motion signals are detected. Due to the receptive field properties of early motion-detecting neurons, they can detect only the component of motion that is in the direction orthogonal to the local orientation of the moving contour (Adelson & Movshon, 1982; Fennema & Thompson, 1979; Marr, 1982; Nakayama & Silverman, 1988; Wallach, 1935). It has been widely hypothesized that the global motions of objects are constructed though an integration of these component-motion signals (e.g., Adelson & Movshon, 1982; Bonnet, 1981; Burt & Sperling, 1981; Hildreth, 1984, 1990; Weiss & Adelson, 1998, 2000; Watson & Ahumada 1985; Weiss, Simoncelli, & Adelson, 2002).

As illustrated in Figure II.13-2a, the detectable component motion for contour moving in a direction that is roughly parallel to its orientation will be of smaller magnitude than that for a contour moving perpendicular to its orientation. It has been shown that such differences in component-motion magnitude correlate strongly with the perceived speed of translating lines (Castet, Lorenceau, Shiffrar, & Bonnet, 1993). For a rotating ellipse, the magnitudes of the component-motion signals scale parametrically with aspect ratio. Specifically, the component motions generated along the contour of a fat ellipse will have smaller magnitudes than those generated along the contour of a skinny ellipse Figure II.13-2b. Furthermore, the degree to which these magnitudes scale with aspect ratio highly correlates with the perceived speed with which ellipses appear to rotate (Caplovitz et al., 2006).



Video II.13-2. Here a set of rotating rectangles will appear to be rotating faster than the following set of rounded rectangles. In fact they are all rotating at the same angular velocity.

A second nonmutually exclusive hypothesis for why the illusion exists is drawn from the fact that the two-dimensional motions of specific form features such as regions of high curvature and contour curvature discontinuities are unambiguous. Unlike component motions signals drawn from regions of low curvature, the motion of such form features can be used to directly determine the global motion of a moving object (Ullman, 1979). There are a number of examples in which the perceived motion of an object more closely follows the motion of its form features than its component motion (Lorenceau & Shiffrar, 1992; Wallach, 1935). There are even stimuli that demonstrate how these two distinct sources of motion information interact with each other (Gori & Hamburger, 2006; Gori, Giora,



Video II.13-3. Here a set of rotating bumps with high and discontinuous contour curvature will appear to be rotating faster than the following set of lower curvature bumps. In fact they are all rotating at the same angular velocity.

Yazdanbakhsh, & Mingolla, 2011; Gori & Yandanbakhsh, 2008; Tse & Hsieh, 2007). As such form features become more distinct, the two-dimensional motion signal they produce will be either more accurate or more heavily weighted in determining the global motion of the object. Because a fat ellipse has relatively low curvature at the ends of its major axis, and thus less distinct form features, it will therefore appear to rotate more slowly than a skinny ellipse.

While the exact mechanism that underlies the spinning ellipse speed illusion remains unknown, for such a simple stimulus, the spinning ellipse has proven to be both aesthetically intriguing and an empirically powerful tool for probing how the visual system constructs motion percepts.

REFERENCES

Adelson, E. H., & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, *30*, 523–525.

Bonnet, C. (1981). Processing configurations of visual motion. In J. Long & A. Bradley (Eds.), *Attention and performance IX*. Hillsdale, NJ: Lawrence Erlbaum.

Burt, P., & Sperling, G. (1981). Time, distance, and feature trade-offs in visual apparent motion. *Psychological Review*, *88*, 171–195.

Caplovitz, G. P., Hsieh, P. J., & Tse, P. U. (2006). Mechanisms underlying the perceived angular velocity of a rigidly rotating object. *Vision Research, 46*(18), 2877–2893.

Caplovitz, G. P., & Tse, P. U. (2007a). Rotating dotted ellipses: Motion perception driven by grouped figural rather than local dot motion signals. *Vision Research, 47*(15), 1979–1991.

Caplovitz, G. P., & Tse, P. U. (2007b). V3A processes contour curvature as a trackable feature for the perception of rotational motion. *Cerebral Cortex*, *17*(5), 1179–1189.

Castet, E., Lorenceau, J., Shiffrar, M., & Bonnet, C. (1993). Perceived speed of moving lines depends on orientation, length, speed and luminance. *Vision Research*, *33*(14), 1921–1936.

Fennema, C., & Thompson, W. (1979). Velocity determination in scenes containing several moving objects. *Computer Graphics and Image Processing*, *9*, 301–305.

Gori, S., Giora, E., Yazdanbakhsh, A., & Mingolla, E. (2011). A new motion illusion based on competition between two kinds of motion processing units: The accordion grating. *Neural Networks*, *24*(10), 1082–1092.

Gori, S., & Hamburger, K. (2006). A new motion illusion: The rotating-tilted-lines illusion. *Perception*, *35*(6), 853–857.

Gori, S., & Yazdanbakhsh, A. (2008). The riddle of the rotating-tilted-lines illusion. *Perception, 37*(4), 631–635.

Hildreth, E. C. (1984). The computation of the velocity field. *Proc R Soc Lond B Biol Sci*, *221*(1223), 189–220; Review.

Hildreth, E. C. (1990). The neural computation of the velocity field. *Res Publ Assoc Res Nerv Ment Dis*, *67*, 139–164; Review.

Kohler, P. J., Caplovitz, G. P., Hsieh, P. J., Sun, J., & Tse, P. U. (2010). Motion fading is driven by perceived, not actual angular velocity. *Vision Research*, *50*(11), 1086–1094.

Lorenceau, J., & Shiffrar, M. (1992). The influence of terminators on motion integration across space. *Vision Research*, *32*(2), 263–273.

Marr, D. (1982). *Vision*. New York: Freeman.

Nakayama, K., & Silverman, G. H. (1988). The aperture problem: II. Spatial integration of velocity information along contours. *Vision Research*, *28*(6), 747–753.

Porter, K. B., Caplovitz, G. P., Kohler, P. J., Ackerman, C. M., & Tse, P. U. (2011). Rotational and translational motion interact independently with form. *Vision Research*, *51*(23–24), 2478–2487.

Stumpf, P. (1911). Uber die Abhangigkeit der visuellen Bewegungsempflndung und ihres negativen Nachbildes von den Reizvorgangen auf der Netzhaut. *Zeitschrift fur Psychologie*, *59*, 321–330.

Stumpf, P. (1996). A gem from the past: Pleikart Stumpf's (1911) anticipation of the aperture problem, Reichardt detectors, and perceived motion loss at equiluminance. *Perception*, *25*(10), 1237–1242; discussion 1235–1237.

Tse, P. U., & Hsieh, P. J. (2007). Component and intrinsic motion integrate in "dancing bar" illusion. *Biological Cybernetics*, *96*(1), 1–8.

Ullman, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.

Wallach, H. (1935). Uber visuell Wahrgenommene Bewegungrichtung. *Psycholgishe Forschung*, *20*, 325–380.

Watson, A. B., & Ahumada, A. J. Jr. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A*, *2*, 232–342.

Weiss, Y., & Adelson, E. H. (1998). Slow and smooth: A Bayesian theory for the combination of local motion signals in human vision. MIT AI Memo 1624 (CBCL Paper 158).

Weiss, Y., & Adelson, E. H. (2000). Adventures with gelatinous ellipses—constraints on models of human motion analysis. *Perception*, *29*, 543–566.

Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, *5*(6), 598–604.

Chapter 14
The Ames Window Illusion and Its Variations

Marcel de Heer and Thomas V. Papathomas

INTRODUCTION

The Ames window (Ames, 1951; Gibson, 1979; Gregory, 2009; Hochberg, 1986; Ittelson, 1952) is an important member of the class of *depth inversion illusions*, for which a three-dimensional (3D) object is perceived in reverse depth; namely, far points on the object appear to be closer than near points on the same object. Other articles on perspective illusions and the hollow-mask illusion in this compendium present other examples of depth inversion illusions. As the name suggests, this illusion was invented by Adelbert Ames, a lawyer and artist who developed about 30 perceptual demonstrations, some of which have become standard material for introductory courses in psychology and for hands-on exhibits in science museums (Behrens, 1999). These demonstrations were accompanied by experiments that were conducted to arrive at reasonable values for the parameters. Among his best-known demonstrations are the Ames room (Dorward & Day, 1997; Gregory, 1966; Rogers, 2010; Runeson, 1988) and the Ames chair (Behrens, 1999; Gregory, 2007). Hochberg (1986) devised variations of the Ames window and argued about a top-down explanation of the illusion, in contrast to Gibson (1979), who hypothesized that the illusion occurs only during weak motion signals.

The fundamental element in the Ames window illusion is shown schematically in Figure II.14-1a; a more realistically painted window, including a shadow, is shown in Figure II.14-1b. Parenthetically, the original illusion included a long, narrow cylindrical horizontal tube that pierced the window plane at a 45° angle and a small cube that was appended near the top of the short, vertical edge of the trapezoid window; interested readers are referred to the original article (Ames, 1951) for more details and a clear description of the observed phenomena. Figure II.14-1a depicts a window frame that is painted on a frontoparallel flat surface—namely, a surface that is normal to the line of sight—using linear perspective, resulting in a physical trapezoid TUVW. Please ignore the black arrow GF for the present; it is used to represent the trapezoid schematically in Figure II.14-3. Even though corners T, U, V and W lie on a frontoparallel plane, observers who view the trapezoid from an adequately long distance perceive the trapezoid as a slanted rectangle, with the long side UT appearing to be closer than the short side VW. More important, even if the window is presented at a slanted orientation such that point U *is* physically further than point V, point U still *appears to be* closer than V, due to linear perspective. We call such an orientation an *illusory-viewed* orientation because the percept has a 3D structure in which depth is reversed with respect to the physical 3D structure. On the other hand, if the trapezoid is oriented with point U closer than point V, then the depth ordering is veridical; namely, point U is perceived to be closer than point V, even though the slant perceived by viewers is not the veridical physical slant. Nevertheless, we call such an orientation a *veridically viewed* orientation, because it does not involve a depth inversion.

The most common version of the Ames window, one that has been used in numerous laboratory and classroom demonstrations, is the rotating version, where the window is set in motion by rotating constantly around a vertical axis PQ at a steady speed, say in the clockwise (CW) direction as

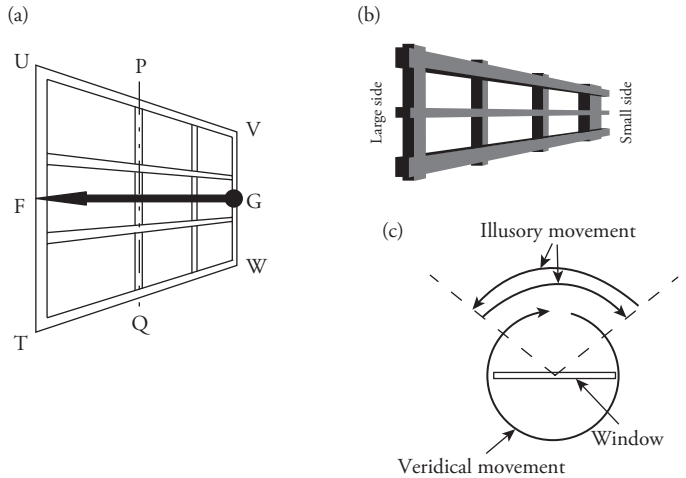
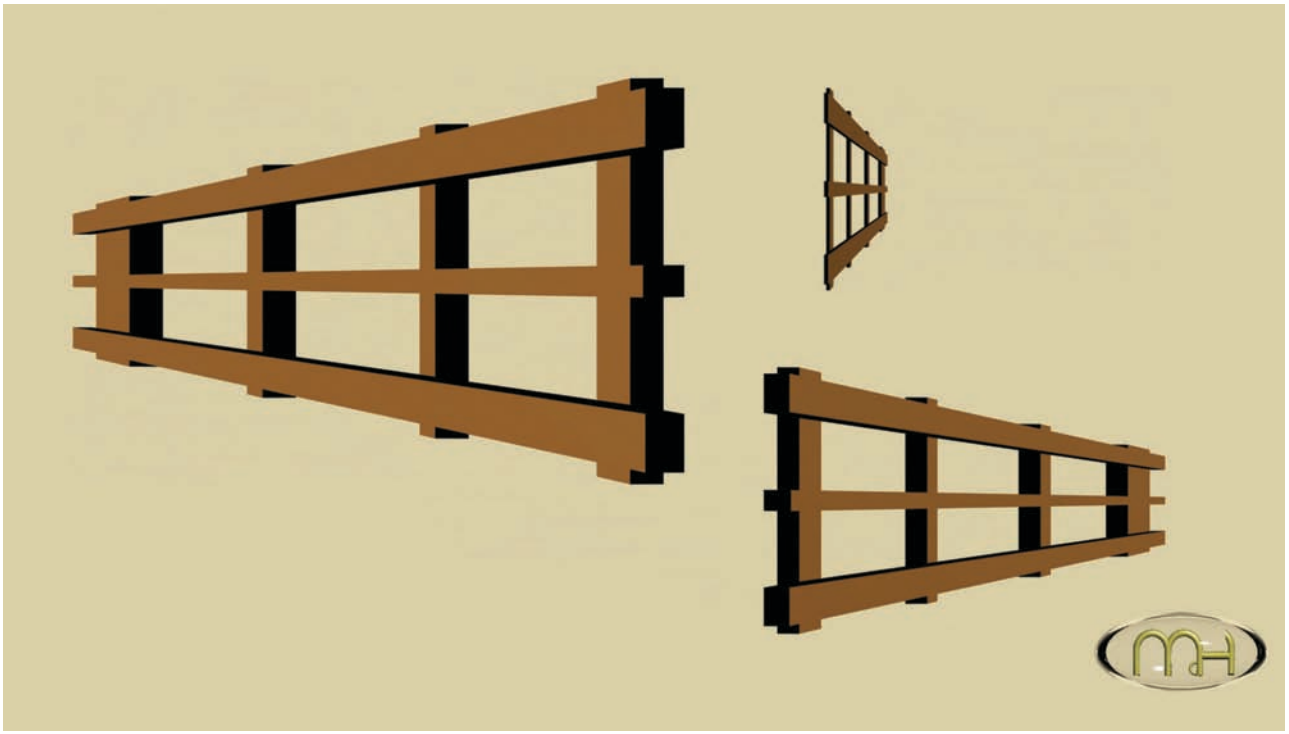


Figure II.14-1. (a) Front view of Ames window lying on a frontoparallel plane. PQ is the axis of rotation. The arrow GF is used to represent the window for the top view of Figure II.14-3, with the filled circle and the arrowhead denoting the short and the long sides of the trapezoid, respectively. (b) A more realistic Ames window, including shadows. (c) Veridical and illusory motion.



Video II.14-1. Realistic Ames windows including shadows.

seen from above. An observer who sees the rotating window from the front perceives the window to oscillate back and forth, as in Figure II.14-1c. There are two basic perceptual modes: (a) initially, when the window turns, starting from its initial frontoparallel position of Figure II.14-1a, it spends the first half-period in illusory-viewed orientations, until it reaches the other frontoparallel position, with side UT to the right of side VW (arrow GF facing to the right), thus during this half-period viewers perceive the window rotating in the *opposite* (counterclockwise [CCW]) direction; (b) during the next half-period, the trapezoid assumes veridically viewed orientations and is perceived to rotate in the physical (CW) direction until it reaches the frontoparallel position shown in Figure II.14-1a, to begin the next period. Overall, instead of seeing the window rotating CW, as it does physically, observers perceive the window oscillating back and forth, alternating between CW and CCW rotation.

So far we have considered a rotating window and a stationary observer. If we reverse the roles and allow an observer to move in front of a static Ames window, we obtain similar phenomena. Namely, the trapezoid will be perceived as a rectangle such that (a) if the window is in an illusory-viewed orientation, the observer will perceive it as moving with him or her; or (b) if the window is in a veridically viewed orientation, then there is no illusion and the observer will perceive it as stationary.

The Ames window presents a challenge to our visual system. In fact, it is deliberately constructed to elicit an illusion by possessing some type of "built-in" perspective. The Ames window can be perceived either as a frontoparallel trapezoid or as a rectangle at an angle. The visual system must decide in favor of the best unambiguous guess between these two competing percepts, as it can have only one stable percept at any given time (Gregory, 1997). When we actually perceive the Ames window illusion, it is because our visual system made the wrong choice. One possible reason is that, in the real world, there are a lot of rectangles (doors, windows, walls, etc.) that are often seen at an angle, and thus we are biased to perceive a trapezoid as a rectangle viewed at an angle. When an Ames window is perceived as a rectangle instead of a trapezoid, it also inherits the characteristics of a rectangle. More on the comparison between quadrangles and trapezoids is in Ames's (1951) original article.

To get a general impression of the Ames-window illision see Video II.14-1.

AMES WINDOWS AND PERSPECTIVE—TOWARD AN EXPLANATION

Ames windows and perspective are closely related to each other. The classical explanation, as outlined previously, involves linear perspective but, as illustrated in Figure II.14-2, straight lines are not necessary to obtain the Ames window illusion. Informal experiments confirm that the de Heer circle of Figure II.14-2a is very effective in eliciting the Ames window illusion; see also the corresponding Video II.14-2 on the DVD that accompanies the compendium. Despite the absence of straight lines in the de Heer circle, it shares a common property with the classical Ames window: they both suggest a shape distortion based on perspective. Notice how the circle appears to be thicker and taller in one end, suggesting that this end is located closer to the viewer.

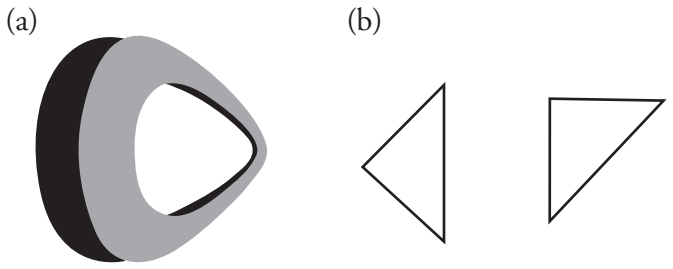


Figure II.14-2. Using different shapes for the Ames window illusion. (a) The de Heer circle. (b) A triangular shaped that was used in the "windsurfer illusion" (Sperling et al. 2009).

In addition, the so-called windsurfer illusion (Sperling, Gobell, & Tseng, 2009) is an Ames window–like illusion with a triangular instead of a trapezoid shape (see Video II.14-3 on the DVD that accompanies the compendium). The fact that these shapes elicit motion illusions that are comparable to the trapezoid Ames window illusion raises the question of whether these illusions share a common underlying mechanism or are caused by different mechanisms. Along the same lines, it would be interesting to transform other common real-world flat objects with a linear perspective or an anamorphosis transformation and test if they exhibit the Ames window illusion, as presented



Video II.14-2. The de Heer circle.



Video II.14-3. Triangular shaped windows used in windsurfer illusion.

later. More systematic experiments are needed to explore these questions.

EXPLANATION FOR THE ILLUSORY MOTION IN THE CLASSIC AMES WINDOW

Ames (1951) himself advanced an explanation for the illusory motion of his window. His explanation uses clever arguments, and it is essentially based on the human bias to perceive trapezoids as slanted rectangles, which is based on prior experience with rectangles. One possible explanation that has been proposed by Papathomas (2007) to account for a wide variety of depth inversion illusions phenomena is equally applicable to the Ames window illusion and has a lot in common with Ames's original explanation. Figure II.14-3 contains top views with the window shown schematically by a thick arrow GF whose circular end G denotes the short vertical base VW and the arrowhead F denotes the long base UT of Figure II.14-1.

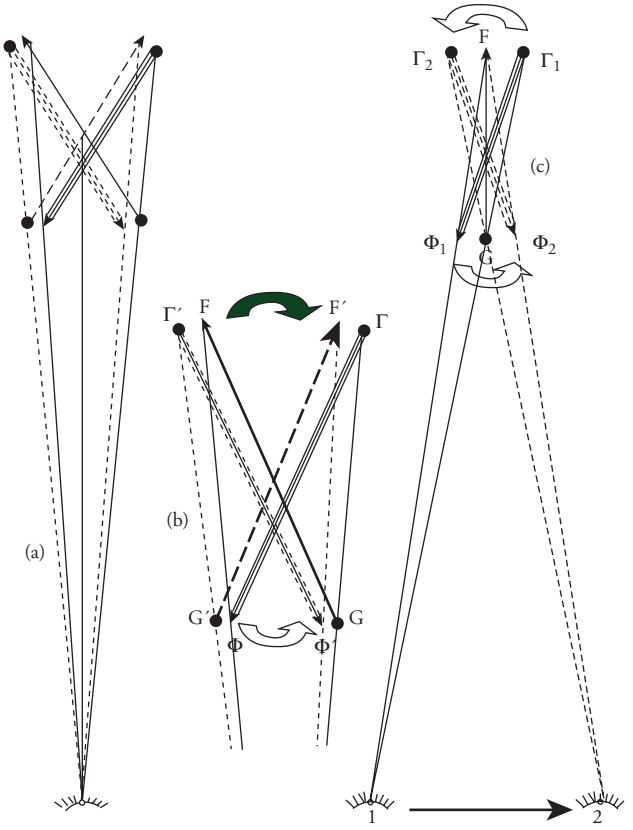


Figure II.14-3. Top views showing the viewer's eye and the trapezoid, represented by the arrow GF; see legend for Figure II.14-1. (a) The trapezoid rotating clockwise. (b) An expanded view of the top part of Figure II.14-3a. Original position GF (solid thick arrow), perceived in inverted depth as ΦΓ (solid multiple-line arrow); final position G'F' (dashed thick arrow), perceived in inverted depth as Φ'Γ' (dashed multiple-line arrow). The physical window rotates clockwise (solid arrow) but the [misperceived] window appears to rotate counterclockwise (open arrow). (c) Window GF is stationary and the eye moves from position 1 to 2 (solid arrow). In positions 1 and 2, the window GF is seen in reverse depth as $\Gamma_1\Phi_1$ and $\Gamma_2\Phi_2$, respectively, eliciting illusory motion (open arrows).

Figure II.14-3a illustrates the case for a stationary observer and a window GF rotating CW, as shown by the solid curved arrow in Figure II.14-3b, which shows the magnified details; in particular, the window starts at position GF (solid thick arrow) and rotates CW to position G'F' (dashed thick arrow). Both positions, however, involve illusory-viewed orientations. Thus instead of perceiving F (edge UT) further away than G (edge VW), we perceive F at point Φ (along F's line of sight), closer than we perceive point G at Γ (along G's line of sight). The resulting percept is shown by the solid multiple-line arrow ΦΓ. Similarly, the terminal position window F'G' (solid thick dashed arrow) is perceived as the depth-inverted window Φ'Γ' (dashed multiple-line arrow). However, as indicated by the open curved arrow, this depth inversion also implies a perceived rotation CCW, opposite to the physical rotation. Notice that this model, appropriately expanded, can account for the shape distortions that are common in this illusion.

Figure II.14-3c illustrates the case for a viewer moving in front of a stationary window oriented along the line of sight of the cyclopean eye. The viewer's eye moves from position 1 (solid lines of sight) to position 2 (dashed lines of sight), as shown by the thick solid arrow, where both positions involve illusory-viewed orientations. Because of the bias to perceive rectangles, the trapezoid GF is perceived as an inverted-depth rectangle $\Phi_1\Gamma_1$ and $\Phi_2\Gamma_2$ for eye positions 1 and 2, respectively. As shown by the open arrows, the perceived rectangle appears to "follow" the viewer, as indeed observed empirically (Ames, 1951).

OBSERVATIONS CONCERNING THE AMES WINDOW AND COGNITION

One of the observations that is confirmed experimentally is that Ames windows without painted "shadows" have a weaker tendency to elicit the Ames window illusion than the ones containing shadows (Ames, 1951; Braden, 1978); to establish a notation, we call these generic classes "weak" and "strong" Ames windows, respectively. The second observation concerns learning. Practicing with Ames windows can change the perception of Ames windows seen afterward. Specifically, some viewers do not perceive the illusion with weak Ames windows. When they later see a strong Ames window, most will perceive the illusion. Subsequently, exposure to the same weak Ames windows that did not initially elicit the illusion will now elicit it more frequently. A third observation is that suggestion and volition play a role on how one perceives the Ames window. Viewers can obtain specific percepts—that they could not attain before—following suggestions by skilled experimenters and practicing to achieve the suggested percepts. In particular, when we showed weak Ames windows, some viewers reported seeing the illusion with the characteristic oscillating back-and-forth motion, but they also reported that, with effort, they perceived the window rotating by 360°.

VARIANTS OF THE AMES WINDOW ILLUSION

The Ames windows in this section are painted the same way on both sides, so that viewers cannot distinguish between the front and the back side. The windows are

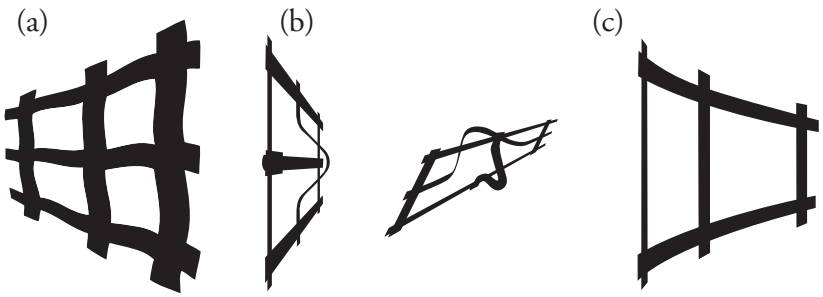


Figure II.14-4. Ames windows containing convex/concave parts. (a) Ames window with "wavy" parts. (b) Ames window with convex/concave distortions along the vertical and horizontal mullions. (c) A curved Ames window.

rendered infinitesimally thin without shadows, using computer graphics techniques. They typically rotate around an axis near the middle of the window; occasionally, they rotate around an axis that is closer to the long or short side. See the videos on the accompanying DVD.

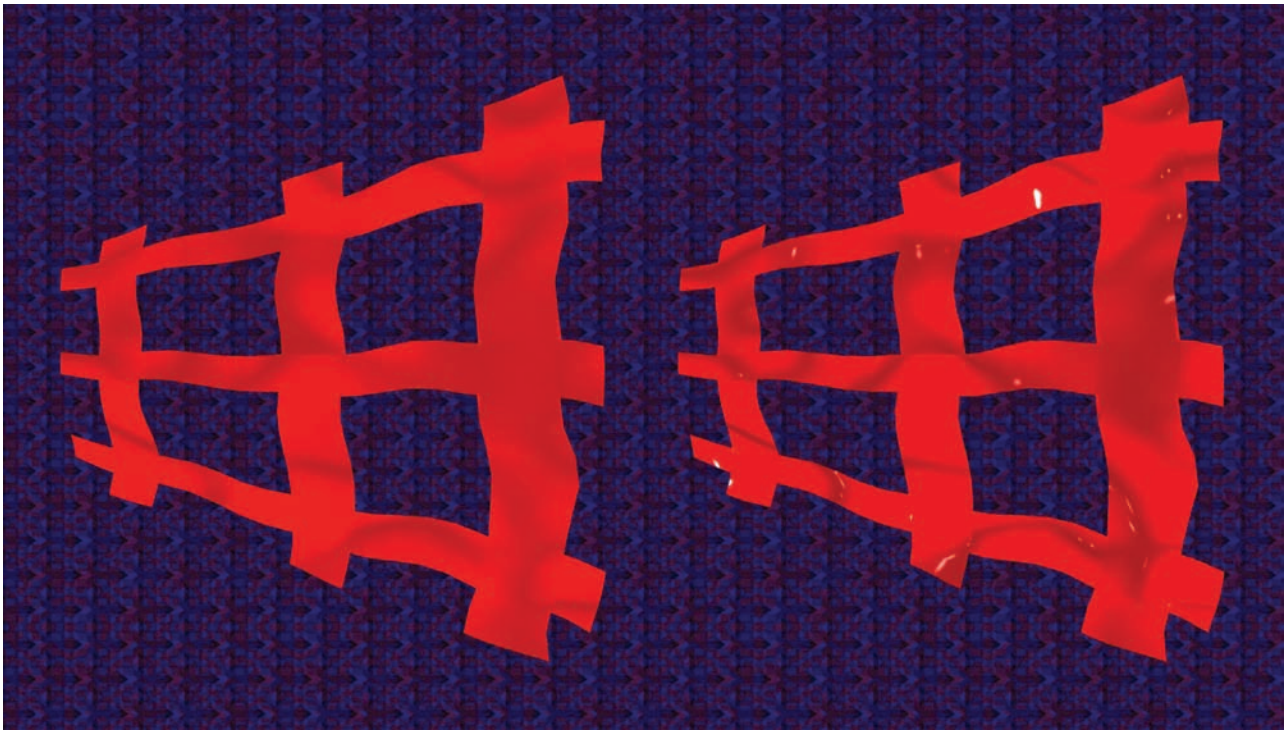
Ames Windows With Convex/Concave Parts

When we perceive an Ames window in the illusory half-period, the depth is inverted not only along the line of sight but also along the dimensions that are orthogonal to the line of sight. Thus when an Ames window has convex and/or concave parts and it transits from the veridical half-period to the illusory half-period or vice versa, the convex and concave parts reverse polarity from convex and concave to concave and convex, respectively. Three examples of Ames windows with convex/concave parts are shown in Figure II.14-4. The Ames window in Figure II.14-4a contains "wavy" parts, generated by introducing distortions along the horizontal and vertical axes. The Ames "fence" in Figure II.14-4b contains convex/concave distortions along both the vertical and horizontal middle bars. When viewers obtain the oscillating illusion in Figure II.14-4b, they obtain a paradoxical percept that one bar of the fence pierces the other bar, which is physically impossible.

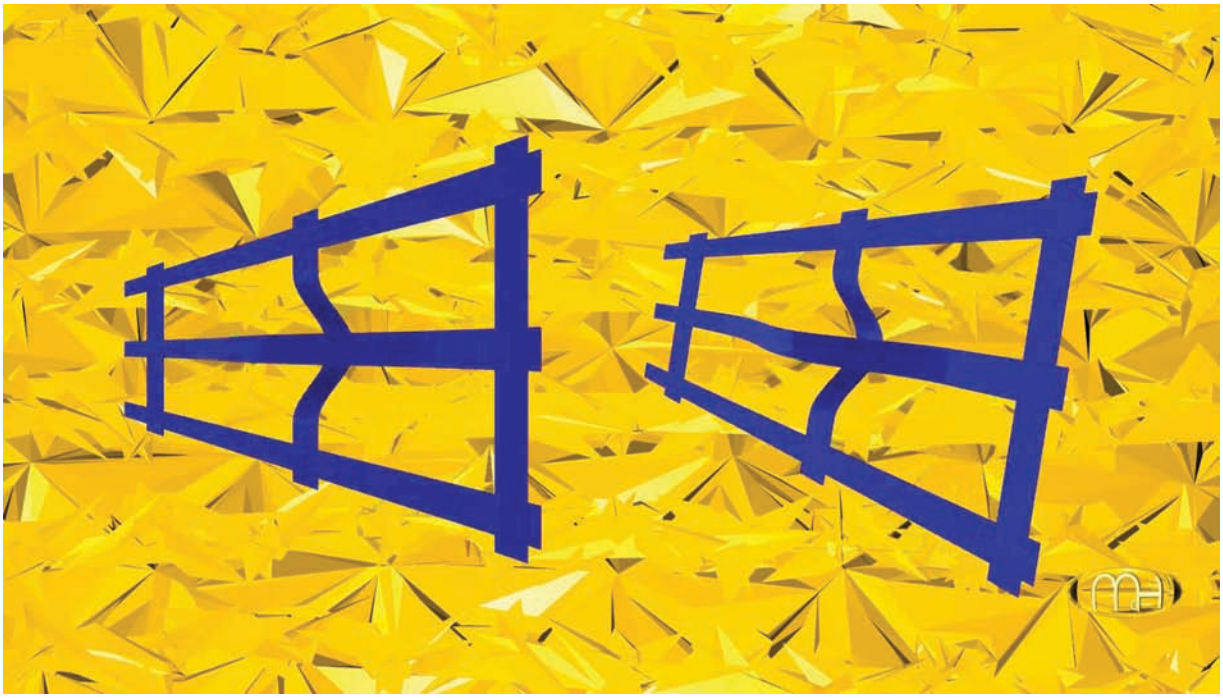
It is also possible to make the entire window convex or concave, as shown in Figure II.14-4c. The accompanying Videos II.14-4, II.14-5, and II.14-6 create very robust percepts, in which convexities are transformed to concavities, and vice versa, as the windows rotate.

Combining Ames Windows: The Stingray and the Butterfly Illusions

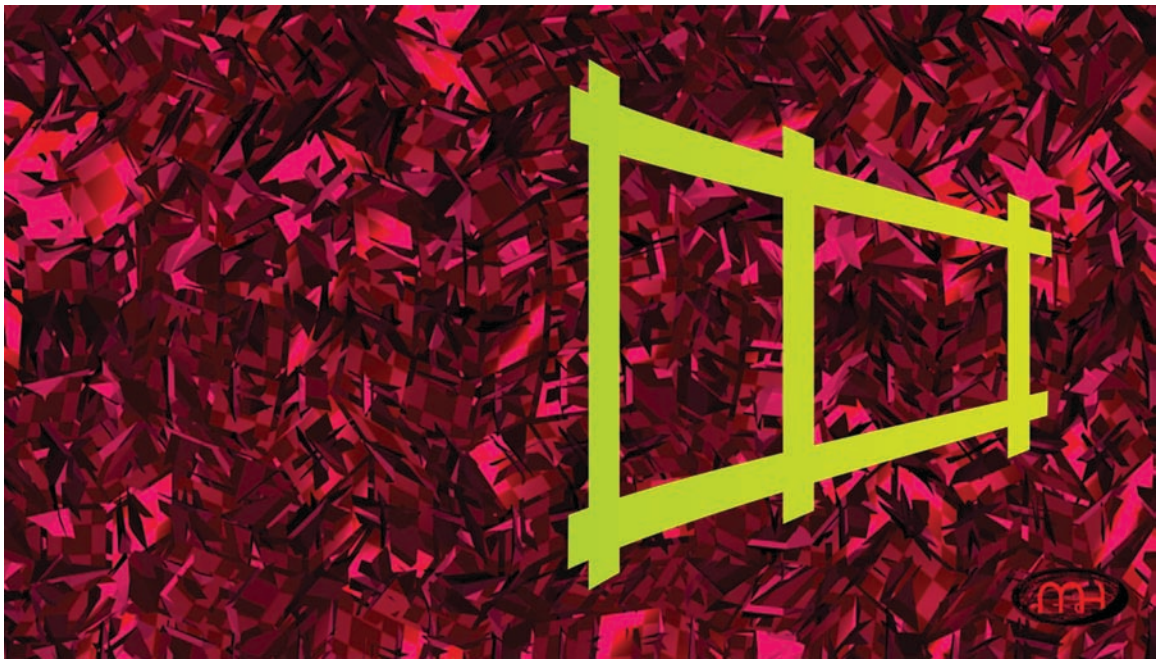
It is possible to combine two or more Ames windows into a new illusory object. An example of this is the de Heer "stingray illusion" that consists of two bent surfaces. Three frames from the stingray Video II.14-7 on the DVD that



Video II.14-4. Ames window with crumpled parts.



Video II.14-5. Ames window with convex concave distortions along the vertical and horizontal mullions.



Video II.14-6. Curved Ames window.

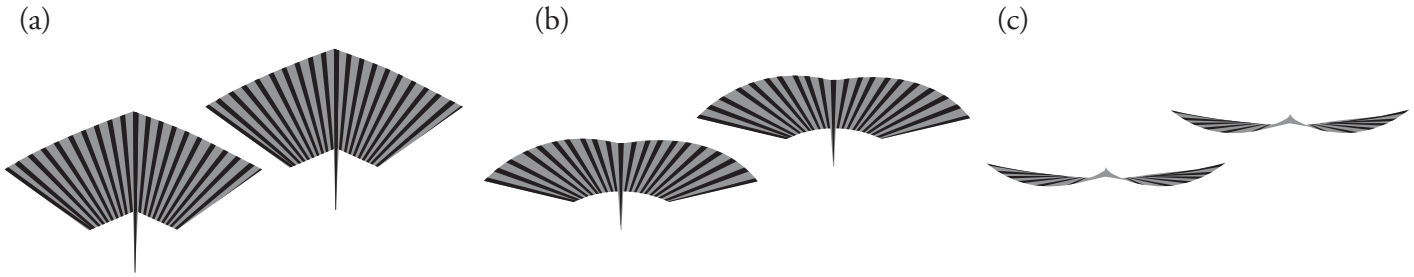
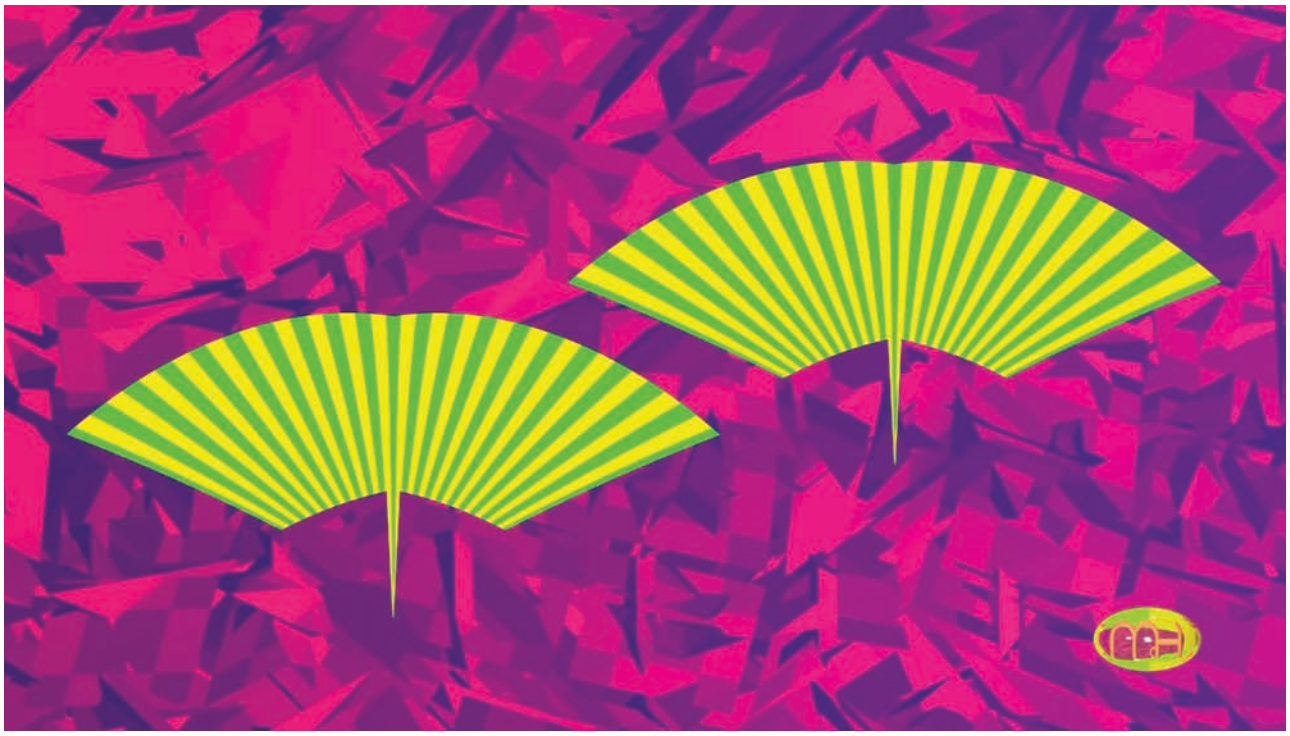


Figure II.14-5. De Heer's "stingray illusion." (a) Frontal view: the curved surfaces are not visible in this view because the stingrays are aligned vertically. (b) View from the front but slightly above the horizon: the curves are now obvious. (c) Top view, showing the vertical alignment.



Video II.14-7. The stingray illusion.

accompanies the compendium are shown in Figure II.14-5. All the Ames windows in the previous sections rotate around a vertical axis. The stingrays are an exception, as they rotate 360° around a horizontal axis, but they appear to oscillate, as if they are swimming toward the viewer. One of the stingrays rotates CW and the other rotates CCW, as seen from the side; at the same time, they are 180° out of phase, such that when one has its long side toward the viewer, the other one has its short side toward the viewer. Of course if we use orthographic projection, we obtain identical images for the two stingrays, even though the stingrays have a different position in space and they move differently. If people see these stingrays rotating in the same direction, they in fact perceive one of them as illusory. Most people see these stingrays "swimming" toward the observer because of the Ames window illusion.

Figure II.14-6 shows another way to combine two Ames "windows" along their short sides to create a single planar surface. When we rotate the combined object around an axis at the junction of the two windows, we get the butterfly illusion. This variation can be appreciated fully when one views the corresponding Video II.14-8 on the DVD that accompanies the compendium, which gives the impression of a butterfly flapping its wings.

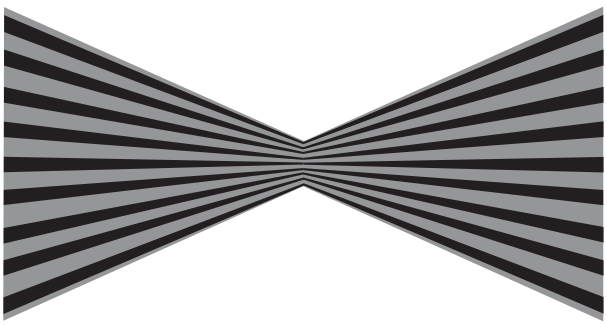


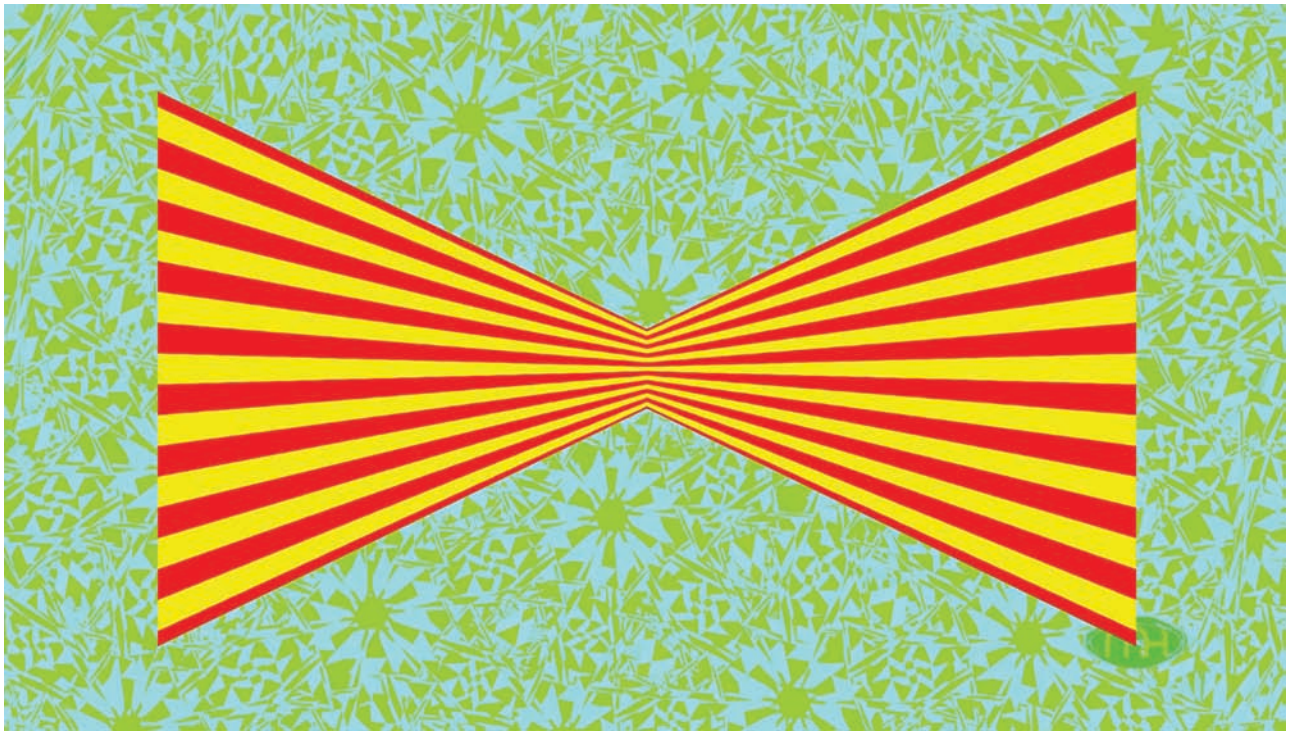
Figure II.14-6. Two Ames windows create the butterfly illusion.

Combining 3D Objects With an Ames Window

When Ames windows are combined with 3D objects, they produce beautiful, paradoxical illusions. Ames windows are usually shown with a stick through them, as Ames (1951) himself did in his experiments. Figure II.14-7 shows an extreme example, where a humanlike 3D figure is embedded in an Ames "fence". The humanlike figure is perceived to rotate normally; the Ames fence, on the contrary, appears to oscillate back and forth. This combination of motions produces a striking illusion: the humanlike figure appears to move in strange ways and mysteriously goes through the fence, as experienced in the Video II.14-9 on the DVD that comes with the compendium.

Illustrating the Ames Window Illusion With a Static Image

Figure II.14-8a and the corresponding Video II.14-10 on the DVD that accompanies the compendium, present two "Ames window–like" surfaces that appear to be mirror-symmetrical versions of each other (with a vertical mirror located between them, orthogonal to the frontoparallel plane). Both curved surfaces may appear to be oriented in space with their long vertical edges closer to the viewer, and each surface appears to have a bump in its middle, akin to



Video II.14-8. The butterfly illusion.

a strange sombrero. However, Figure II.14-8a presents a remarkable trick that illustrates *statically* the bistability of the Ames window illusion. The actual geometry of the sombrero in the right panel of Figure II.14-8a is not the one described previously. In fact, it is very different, as illustrated in the right panel of the top view in Figure II.14-8b: it consists of a surface with a "dimple," rather than a bump, with its long edge away from the viewer.

In fact, one may perceive each sombrero in two stable states, and we illustrate this for the sombrero on the right-hand panel of Figure II.14-8a: One stable percept is the *illusory* one described previously (long edge closer to viewer and a bump in the middle). If one tries, one can perceive the *veridical* geometry by volitionally "turning" the bump into a concavity; in this case, the hat's orientation in space also switches, with the short vertical edge appearing closer to the viewer (as the right panel in the top view of Figure II.14-8b indicates). Even if the static images of Figure II.14-8a are bistable, they provide motion cues to recover the true depth when they rotate. Nevertheless, viewers perceived a depth-inversion illusion. The sequence of perceptual switches from the illusory to the real geometry can happen in another way: the viewer can achieve a percept change in the orientation of the hat such as to perceive the short vertical edge as closer; then the curved central part changes percept automatically from a bump to a concavity. In sum, the two stable percepts for the sombrero in the right panel of Figure II.14-8a are (a) long edge closer to the viewer and a bump in the middle (illusory) and (b) short edge closer to the viewer and a concavity in the middle (veridical). Thus the perceived surface orientation and the polarity (concavity or convexity) of the central surface feature are yoked. This allows the viewer to perceive the two surfaces in four possible combinations, depending on whether the left and the right sombrero are perceived with their veridical or their illusory percept.

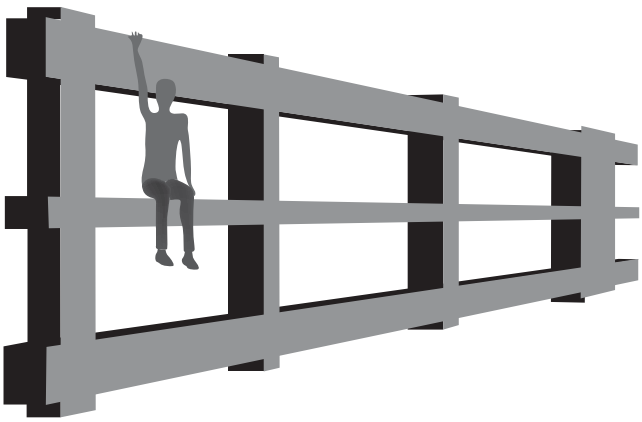
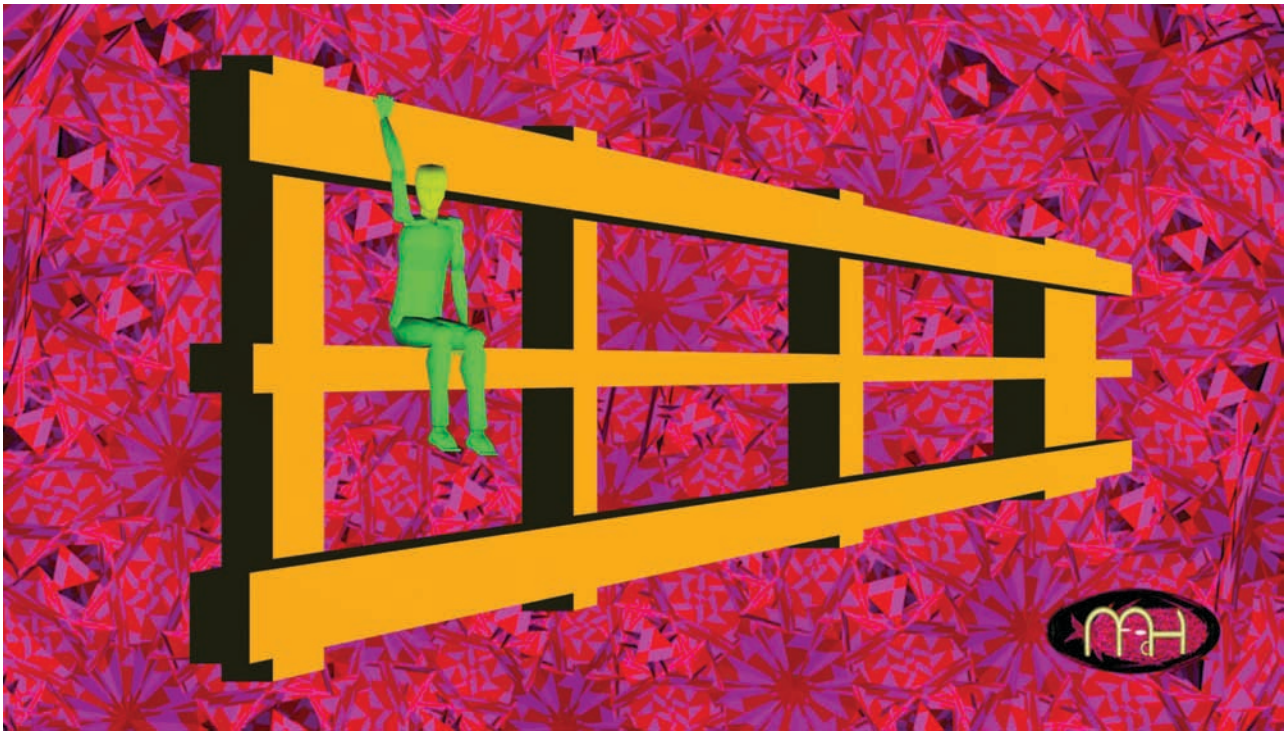


Figure II.14-7. Humanlike figure sitting on an Ames fence.

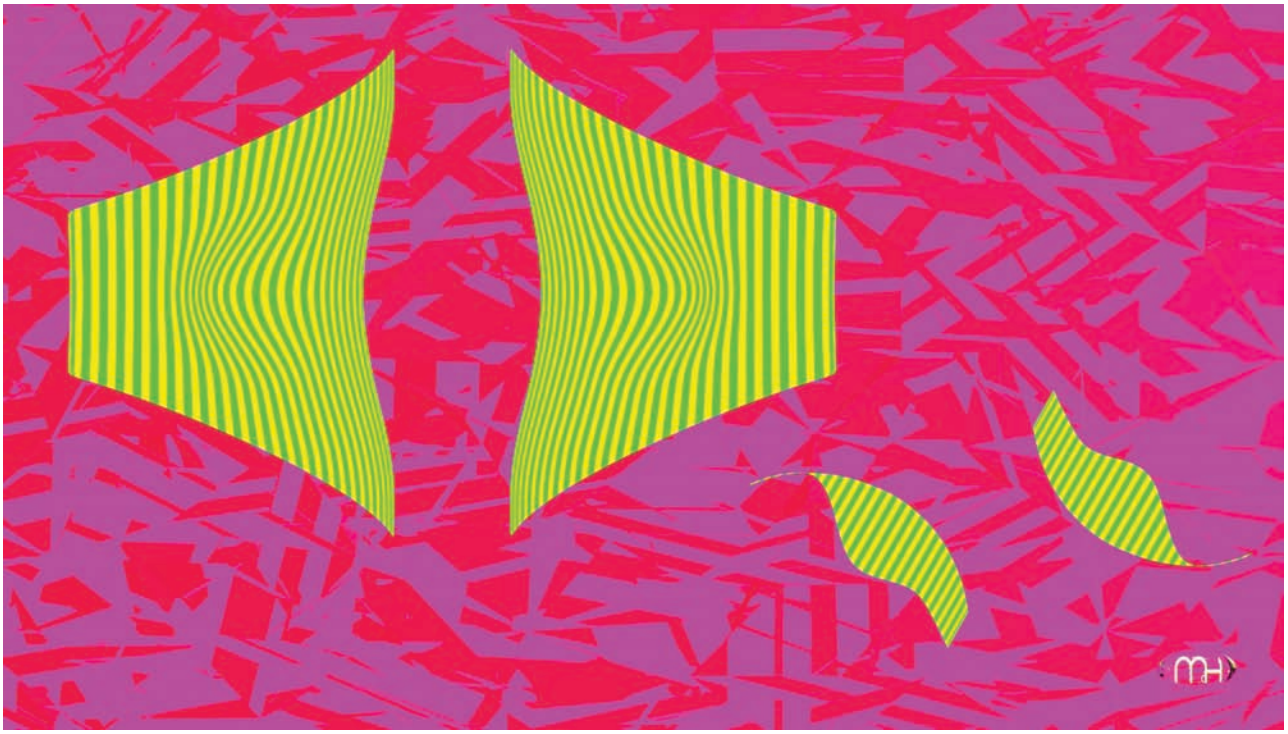
The Video II.14-10 is much more compelling in bringing out this bistability. It must be pointed out that each sombrero does not make a full 360° rotation, but they actually oscillate back and forth, as shown on the top view in the Video II.14-10 Depending on the geometrical interpretation of the two sombreros, we may perceive them oscillating veridically (out of phase, with the long edge of one sombrero approaching and the long edge of the other sombrero receding) or in an illusory fashion (in phase, with the long edges of both sombreros approaching together or receding together).

A Tilted Ames Window

A fascinating version of the Ames window results when the window is tilted, as shown in Figure II.14-9, which



Video II.14-9. Humanlike figure sitting on an Ames fence.



Video II.14-10. The de Heer's sombrero illusion.

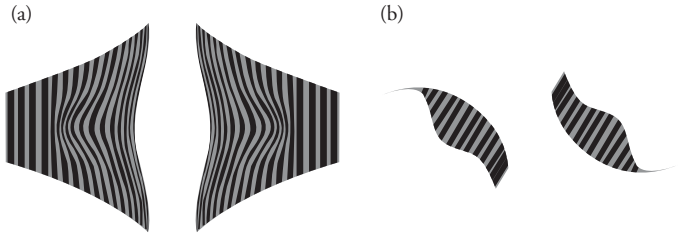


Figure II.14-8. de Heer's "sombrero illusion." Figure (a) shows the front view. Figure (b) shows the top view.

represents a frame in the Video II.14-11 sequence on the compendium's DVD. As the reader can attest, it is difficult to tell what the shape of such a window really is. The window of Figure II.14-9 could be a skewed trapezoid on a frontoparallel plane; it could also be a rectangle in an oblique orientation; in fact, it could even be a quadrilateral from an infinite set of combinations of shapes and orientations. The recovery of the true shape and orientation is, of course, one of the basic problems that face the visual system. We believe that the tendency to perceive such trapezoids as rectangles is one possible explanation for the remarkable transition that occurs in this version of the Ames window illusion.



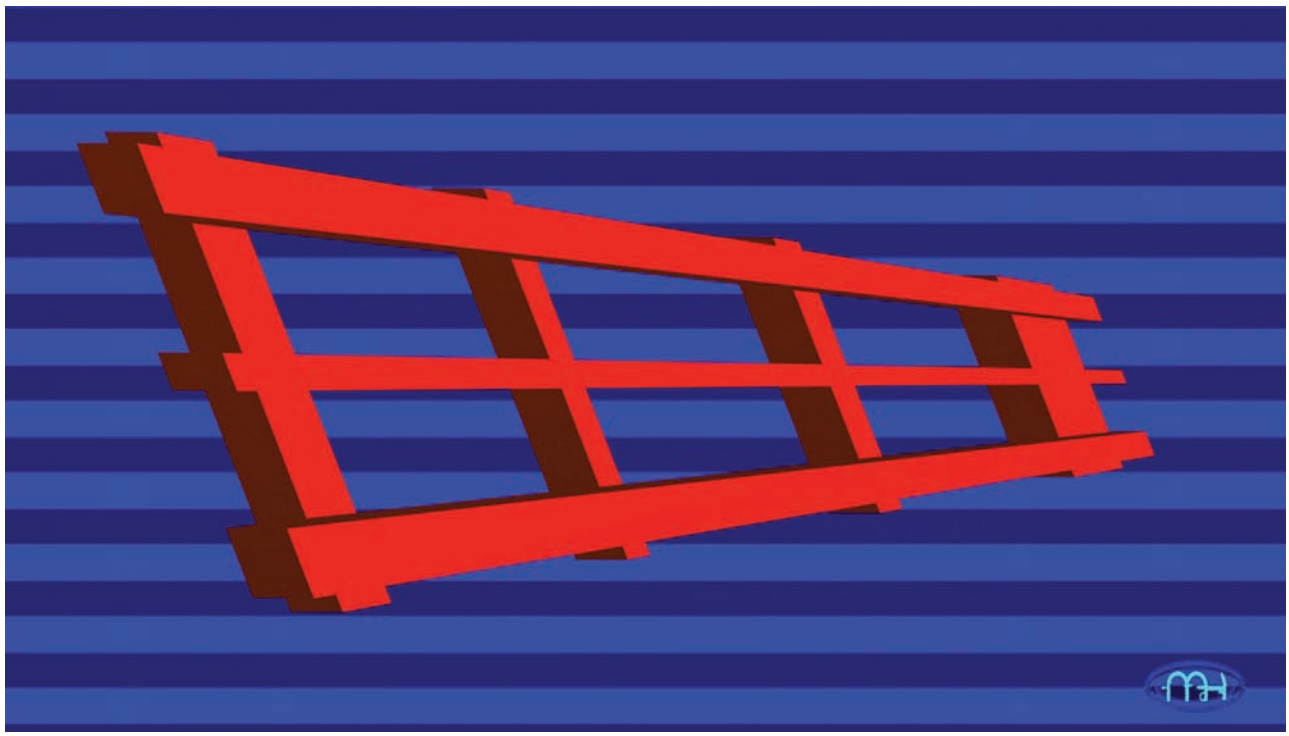
Figure II.14-9. A tilted Ames window.

CONCLUSIONS

The Ames window illusion is a powerful demonstration of the role played by past experience on perceiving objects in the world. In most of the examples in this chapter, past experience with rectangles and perspective can dominate over sensory visual signals that are provided by motion parallax and stereoscopic disparity. Just as in the case of the hollow mask and reverspectives, this illusion can be constructed using physical 3D objects, which offer more natural stimuli than representations of these objects by planar (two-dimensional) images for experimentation, as well as for interactive demonstrations.

ACKNOWLEDGMENTS

The authors wish to thank family members and friends who were almost coerced to experiment with all kinds of Ames windows. We thank Jillian Nguyen for providing us with valuable feedback in preparing this chapter. We also thank Sjef Sprangers, who taught us a lot about the use of 3D software.



Video II.14-11. Tilted Ames window.

REFERENCES

Ames, A. Jr. (1951). Visual perception and the rotating trapezoidal window. *Psychological Monographs: General and Applied*, *65*(7), i–32.

Behrens, R. R. (1999). Adelbert Ames, Fritz Heider and the chair demonstration. *Gestalt Theory*, *21*(3), 184–190.

Braden, J. H. (1978). The effect of varying linear perspective, movement parallax, and speed of rotation on perceived oscillation of two-dimensional trapezoidal stimuli. *Perception & Psychophysics*, *23*(1), 51–57.

Dorward, F. M., & Day, R. H. (1997). Loss of 3D shape constancy in interior spaces: The basis of the Ames-room illusion. *Perception*, *26*(6), 707–718.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton.

Gregory, R. L. (1966). *Eye and brain.* London: Weidenfeld & Nicolson.

Gregory, R. L. (1997). Knowledge in perception and illusion. *Philosophical Transactions of the Royal Society of London, B, Biological Sciences*, *352*, 1121–1128.

Gregory R. L. (2007). Objects and pictures. In F. A. J. L. James (Ed.), *Christmas at the Royal Institution: An anthology of lectures by M. Faraday, J. Tyndall, R. S. Ball, S. P. Thompson, E. R. Lankester, W. H. Bragg, W. L. Bragg, R. L. Gregory, and I. Stewart* (p. 315). London: World Scientific.

Gregory, R. L. (2009). Looking through the Ames window. *Perception*, *38*, 1739–1740.

Hochberg, J. (1986). Visual perception of real and represented objects and events. In N. J. Smelser & D. R. Gerstein (Eds.), *Behavioral and social science: Fifty years of discovery. In commemoration of the fiftieth anniversary of the "Ogburn Report": Recent social trends in the United States* (pp. 249–298). Washington, DC: National Academy Press.

Ittelson, W. H. (1952). *Ames demonstrations in perception: A guide to their construction and use.* Princeton, NJ: Princeton University Press.

Papathomas T. V. (2007). Art pieces that "move" in our minds—An explanation of illusory motion based on depth reversal. *Spatial Vision*, *21*, 79–95.

Rogers, B. (2010). Stimuli, information, and the concept of illusion. *Perception*, *39*, 285–288.

Runeson, S. (1988). The distorted room illusion, equivalent configurations, and the specificity of static optic arrays. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 295–304.

Sperling, G., Gobell, J., & Tseng, C.-H. (2009). Trapezoidal illusions: Windsurfers versus runways [Abstract]. *Journal of Vision*, *9*(8), 691.

Chapter 15
Three-Dimensional Müller-Lyer Illusion

Theoretical and Practical Implications

Patricia R. DeLucia

INTRODUCTION

The three-dimensional (3D) Müller-Lyer illusion is an embodiment of the two-dimensional (2D) Müller-Lyer illusion in a solid construction. Figure II.15-1 depicts the Müller-Lyer pattern in two dimensions (drawn lines). Figure II.15-2 shows a solid lucite Müller-Lyer object (used by DeLucia & Hochberg, 1991). Figure II.15-3 depicts the pattern in three dimensions with familiar objects (books, funnels). In both 2D and 3D stimuli, the apex-in distance (>—<) appears longer than the apex-out distance (<—>), although both are in fact equal. Since its introduction over a century ago (Müller-Lyer, 1889/1981), the Müller-Lyer illusion has incited numerous studies and explanations (Boring, 1942). The occurrence of the Müller-Lyer illusion in 3D constructions has significant theoretical and practical implications.

THEORETICAL IMPLICATIONS

Perspective and Misapplied Constancy Scaling Theories

The 3D Müller-Lyer illusion has particular importance for the time-honored and widely cited perspective and misapplied constancy scaling explanations exemplified by Gregory (e.g., Gregory, 1963; see also Coren & Girgus, 1978; Woodworth, 1938). According to this theory, observers treat the Müller-Lyer pattern as a 2D projection of 3D objects and the fins are misinterpreted as representing the depth cue of linear perspective. Gregory proposed that in accordance with this depth cue, the fins of the apex-out Müller-Lyer figure are interpreted as receding away from the eye analogous to the walls of an external corner of a building. Conversely, the fins of the apex-in Müller-Lyer figure are considered as advancing toward the eye analogous to the walls of an internal corner of a room. The linear perspective cue activates a constancy-scaling mechanism through which the size of the shaft in the apex-in pattern is enlarged to compensate for its presumed farther distance (and shortened in the presumed closer apex-out pattern). Although putatively this scaling mechanism ordinarily results in size constancy (so that objects moving away from the eye do not appear to shrink due to decreasing retinal size), applying this same mechanism to the 2D Müller-Lyer figure is inappropriate because the oblique patterns projected by the fins are not due to changes in distance. Consequently, an illusion occurs. However, the misapplied constancy theory predicts that the Müller-Lyer illusion will not occur with 3D objects because there are no 2D lines to misinterpret (Massaro & Anderson, 1970; Zanforlin, 1967). Therefore, the occurrence of the 3D Müller-Lyer illusion provides evidence that disconfirms the theory.

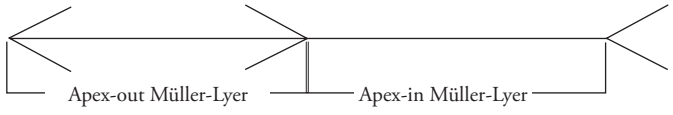


Figure II.15-1. The 2D line-drawn Müller-Lyer illusion (Brentano form; Coren & Girgus, 1978). The apex-in distance appears longer than the apex-out distance although both are equal.

DeLucia and Hochberg (1991) considered two related explanations of the 3D Müller-Lyer illusion. First, they noted that a 3D Müller-Lyer object can present diagonal edges on the retina that are similar to those created by the 2D figure, leading to misinterpretations regarding depth and resurrecting the misapplied constancy mechanism. Second, they considered Gillam's (1978) proposal that observers perceive the fins as rotated closer to 90° than actual because of a scaling process for achieving constancy (i.e., when right angles in 3D space project nonorthogonal angles on the retina) without relying on depth cues. These explanations are difficult to reconcile with the finding that the 3D Müller-Lyer illusion occurred even when a Müller-Lyer object (created with books) was oriented such that it did not present any diagonal edges to the eye (DeLucia & Hochberg, 1991). DeLucia and Hochberg noted that their results did not rule out the averaging or assimilation theories (see Coren & Girgus, 1978), which propose that observers respond to figural characteristics other than the distances they are asked to judge.

Level of Processing

The 3D Müller-Lyer illusion has been used to examine the level of processing at which the illusion occurs. For

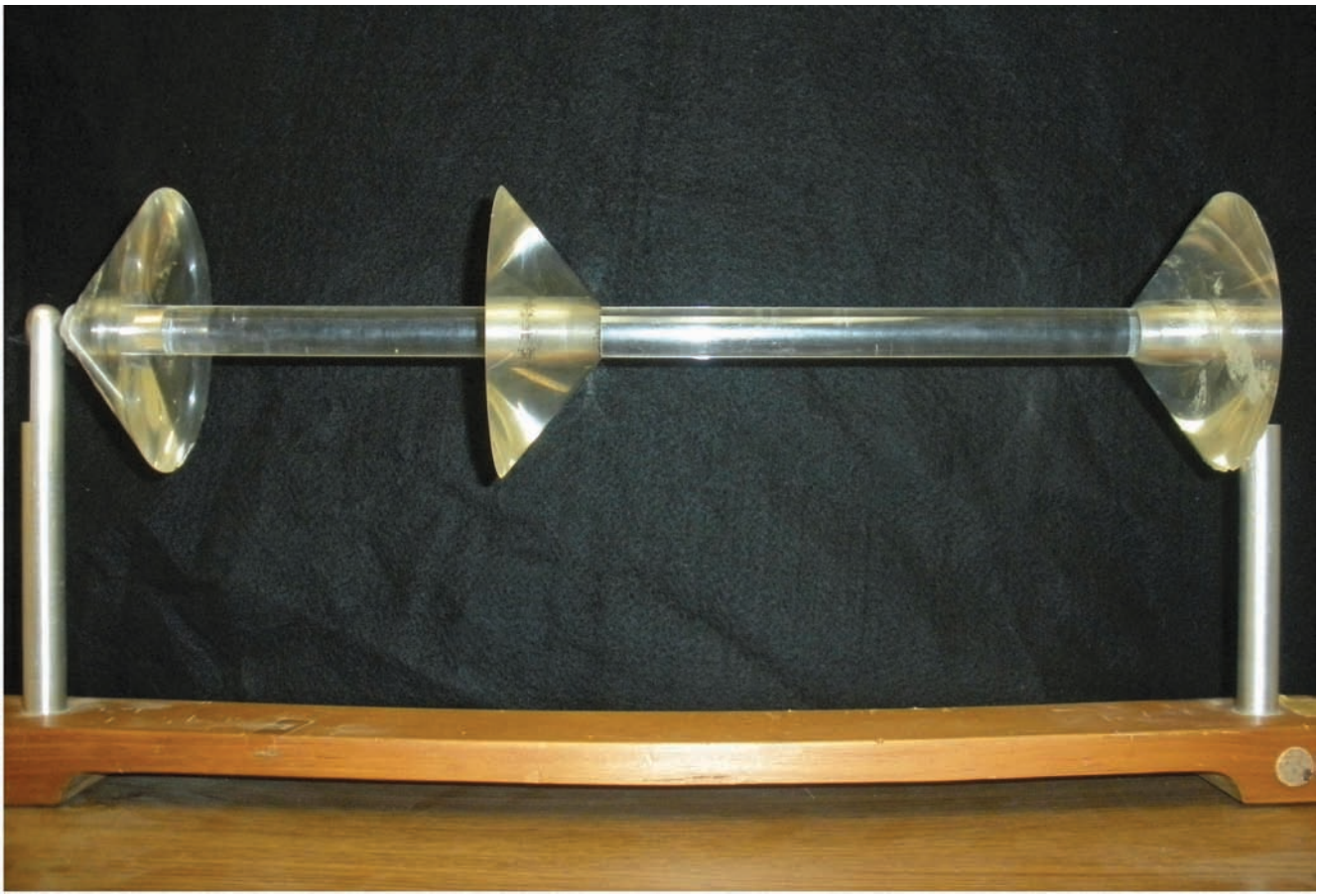


Figure II.15-2. One of the 3D lucite Müller-Lyer objects (50° fin angle) used by DeLucia and Hochberg (1991).

example, Nijhawan (1991, 1995) used 3D Müller-Lyer patterns to examine the retinal image's contribution to the illusion. Rectangular sheets of transparent plexiglass were covered with fluorescent material at the edges to create outlines of the Müller-Lyer pattern or with random dots that created a pattern devoid of lines (and thus depth cues needed for constancy scaling). The illusion occurred in both cases, indicating that a retinal pattern consisting of Müller-Lyer shapes was not needed for the illusion to occur. When the fins were rotated in 3D space (similar to the stereoscopically tilted fins used by Georgeson & Blakemore, 1973) to be either parallel or orthogonal to the shaft's plane, the illusion was smaller when the 3D shape differed from the 2D shape even though both configurations projected the same traditional Müller-Lyer pattern on the retina. Furthermore, when the 3D oblique fins were positioned to project straight rather than oblique lines on the retina, the illusion occurred nevertheless. Nijhawan concluded that the retinal image created by the 2D Müller-Lyer pattern is not necessary or sufficient for the illusion to occur and that the illusion occurs after depth information is processed.

Müller-Lyer figures that appear 3D with the use of stereoscopic displays also have been used to examine the processing stage at which the illusion occurs. The occurrence of the Müller-Lyer illusion with such displays (Georgeson & Blakemore, 1973; Julesz, 1971; Papert, 1961) has been considered evidence that the illusion is due to processes more central than the retina; that is, after stereopsis occurs (Julesz, 1971; Papert, 1961).

Cross-Modality Testing

Three-dimensional Müller-Lyer configurations have been used to determine whether the Müller-Lyer illusion occurs in the haptic modality. For example, a robust illusion occurred when participants used their index finger to trace a Müller-Lyer pattern composed of raised lines (Heller et al., 2002). The haptic illusion did not depend on visual experience: The illusion did not differ among participants who were blind since birth, had lost sight after one year of life, had low vision (perception of patterns was minimal), or were sighted but blindfolded. Moreover, the familiar inverse relationship between fin angle and illusion magnitude observed in the visual modality (Berliner, 1948) also occurred in the haptic modality (using sighted blindfolded participants). Although this finding is consistent with the idea that the visual and haptic Müller-Lyer illusions result from similar mechanisms, different mechanisms can lead to the same outcome (Genatz & Hatwell, 2004; Heller et al., 2002). The haptic illusion has raised doubts about the misapplied constancy scaling explanation (Day, 1965; Gentaz & Hatwell, 2004), but Gregory (1963) argued that results from the haptic modality should not be considered evidence against explanations of the visual illusion (Day, 1965).

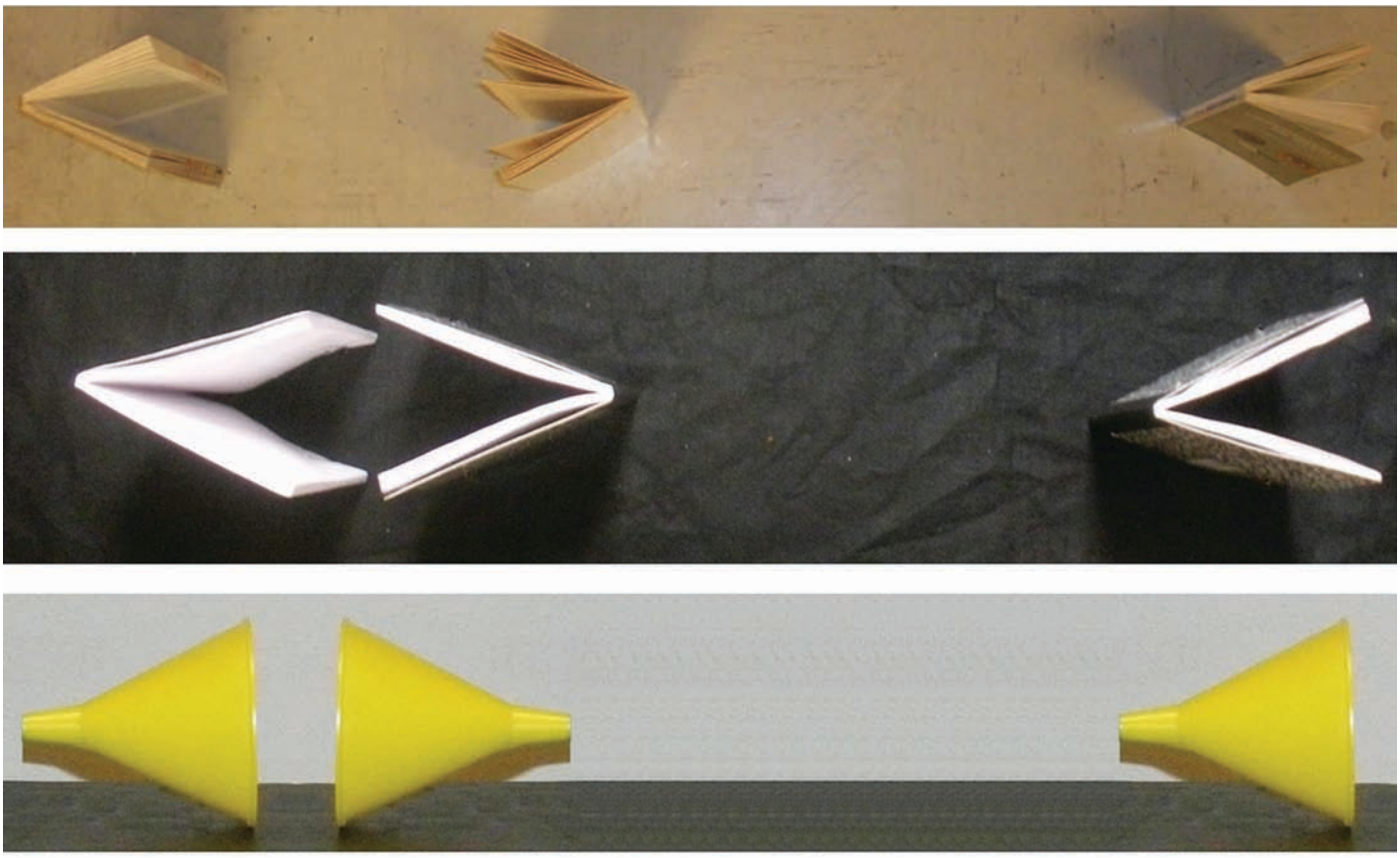


Figure II.15-3. The 3D Müller-Lyer illusion with familiar solid objects (Brentano form; Coren & Girgus, 1978). In 3D space, the horizontal distances between each pair of fins (book spines, funnel vertices) were equal in length. All photographs were modified with Adobe Photoshop to ensure that, in the image, the apex-in and apex-out distances (distances between tops of book spines, vertices of funnels) were equal. Top panel: Photograph of a Müller-Lyer pattern composed of paperback books. Lines due to tile edges were removed with Adobe Photoshop. Middle panel: Photograph of a Müller-Lyer pattern composed of composition notebooks. Paper clips used to hold pages together were removed with Adobe Photoshop. Bottom panel: Müller-Lyer pattern created with a digital photograph of a single plastic funnel and MS-Paint.

PERCEPTION–ACTION DISSOCIATIONS

More recently, the 3D Müller-Lyer illusion has been used to evaluate dissociations between perception and action (for a review, see Bruno & Franz, 2009). Such dissociations are predicted from Goodale and Milner's (1992; Milner & Goodale, 1995) proposal that there are two functionally distinct pathways in the visual system. The ventral pathway (primary visual cortex to inferior temporal cortex) is responsible for visual perception (e.g., recognizing and identifying objects) and relies on object-centered coordinates. The dorsal pathway (primary visual cortex to the posterior parietal cortex) is responsible for visually guided actions (e.g., grasping and intercepting objects) and relies on egocentric coordinates. According to this two visual systems hypothesis, illusions should affect perceptual judgments but should not affect visually guided actions. For example, when participants grasped the wooden shaft of a Müller-Lyer figure, maximum grip aperture did not exhibit an influence of the illusion, consistent with Goodale and Milner's hypothesis (Otto-de Haart, Carey, & Milne, 1999).

Wraga, Creem, and Proffitt (2000) also provided results consistent with the two visual systems hypothesis, with large Müller-Lyer configurations created from ribbons and hoops placed on the floor. Participants exhibited the illusion when they estimated line length verbally and by blindly walking the estimated length after viewing the pattern from a stationary position located 1.5 m from the pattern's endpoint (to elicit object-centered encoding). However, the illusion was not observed with blind walking after the pattern was viewed while standing on an endpoint of the pattern (to elicit egocentric encoding).

However, evidence regarding the two visual systems hypothesis is mixed. For example, there are studies of grasping (e.g., Heath, Rival, Neely, & Krigolson, 2006) and a meta-analytic review of grasping studies (Bruno & Franz, 2009) that contradict the hypothesis. Moreover, the use of illusions to test the two visual systems hypothesis has been criticized on methodological (e.g., Wraga et al., 2000) and logical grounds (DeLucia, Tresilian, & Myer, 2000).

Finally, 3D Müller-Lyer constructions have been used to test the illusion during active viewing. DeLucia and Hochberg's (1991) participants walked back and forth to view a large Müller-Lyer pattern constructed from plywood (Figure II.15-4). The occurrence of the illusion indicated that it is not limited to passive viewing of 2D line drawings: The illusion can occur with moving participants who actively view 3D objects (see movingwoodml.avi for

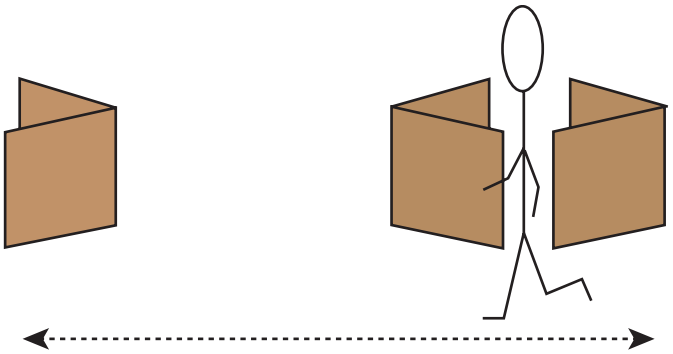


Figure II.15-4. Schematic representation of the 3D Müller-Lyer illusion constructed from large plywood and viewed by actively moving observers. See movingwoodml.avi for a digital recording of a wooden 3D Müller-Lyer pattern viewed through a moving camera.

a digital recording of a wooden 3D Müller-Lyer pattern viewed through a moving camera).

VARIANTS OF THE THREE-DIMENSIONAL MÜLLER-LYER ILLUSION

The 3D Müller-Lyer illusion has been demonstrated with a variety of solid objects. These include cardboard and stick figures (Massaro & Anderson, 1970), wooden mats (Zanforlin, 1967), solid clear cones (DeLucia & Hochberg, 1991), plexiglass panels (Nijhawan, 1991), large plywood constructions viewed by actively moving observers (DeLucia & Hochberg, 1991), and ordinary books (DeLucia & Hochberg, 1991).

Two key parameters that affect the 3D Müller-Lyer illusion are fin angle and practice. Illusion magnitude decreases with increases in fin angle (DeLucia & Hochberg, 1991; Massaro & Anderson, 1970; Nijhawan, 1991) and number of trials (Bruno & Franz, 2009; Massaro & Anderson, 1970). Fin angle has been manipulated to evaluate whether the 2D Müller-Lyer illusion and the 3D Müller-Lyer illusion are the same illusion (DeLucia & Hochberg, 1991) and share a common causal mechanism (Nijhawan, 1991).

Related variations of the 3D Müller-Lyer occur when the Müller-Lyer figure is embedded in 2D images of 3D scenes. Fisher and Lucas (1969) provided numerous examples of illusory figures embedded in line drawings and photographs of familiar objects such as the Müller-Lyer pattern within the pleats of an accordion's bellows and the spines of a book. Moreover, the Müller-Lyer illusion occurred when the pattern existed within real-world maps and affected both distance estimates and decisions based on such estimates (Gillan, Schmidt, & Hanowski, 1999). These examples show that the Müller-Lyer illusion can occur in real-world contexts.

RELATED ILLUSIONS AND PRACTICAL IMPLICATIONS

The 2D geometrical illusions and visual illusions more generally have long been criticized as being restricted to impoverished, artificial stimuli such as line drawings and as having little relevance to the understanding of ordinary perception (Gibson, 1950, 1966, 1979; Michaels & Carello, 1981; Pieron, 1955, cited in Fisher & Lucas, 1969). The occurrence of the 3D Müller-Lyer illusion, particularly in rich environmental settings (i.e., with motion of the observer in textured environments and familiar objects such as books) refutes this criticism. Moreover, any illusion that can occur in 3D contexts has practical importance and is an important concern for visual science.

Furthermore, the Müller-Lyer illusion is not the only geometrical illusion that occurs in three dimensions. The Ponzo illusion was demonstrated with real-world views of railroad tracks (Leibowitz, Brislin, Perlmutter, & Hennessy, 1969) and with large wooden boards (Brislin & Keating, 1976) and small wooden strips (DeLucia & Hochberg, 1991) viewed freely in real-world settings. Similarly, the vertical–horizontal illusion was demonstrated with real-world objects on a university campus (Chapanis & Mankin, 1967). The Hering illusion occurred with steel-wire constructions (DiLollo & Marshall, 1969), and the Oppel-Kundt illusion occurred with light-emitting diodes arranged in 3D space (Deregowski & McGeorge, 2006). Chadwick (1976) provided examples of how a variety of illusions can occur when geologists read rock formations. More generally, it has been recognized for centuries that illusions occur in 3D structures, and architects have corrected for them (Coren & Girgus, 1978; Goodyear, 1899). Examples of illusions in large constructions include the Parthenon of Athens (Coren & Girgus, 1978; Fisher & Lucas, 1969), the Gateway Arch in St. Louis, Missouri (Coren & Girgus, 1978), and Borromini's perspective arcade in Rome's Palazzo Spada (Pirenne, 1970).

It is important to recognize that even if illusions were confined to 2D displays, it is critical to predict their occurrence and size because people rely increasingly on 2D displays (DeLucia & Hochberg, 1991). Examples include digital images (e.g., computer displays), x-rays (e.g., medical diagnostics, baggage screening), camera images (e.g., minimally invasive surgery), and radar (e.g., weather, air-traffic control).

Although it is common to associate illusions with negative effects on behaviors (errors), illusions also can be used to enhance behavior. For example, when transverse lines that decreased in spacing were painted on a road to create an illusion of speeding up, mean driving speed and accidents attributed to speed decreased (Godley, Fildes, & Triggs, 1997). Similarly, when a Wundt illusory pattern was painted on a rural curved road with a high rate of accidents to create the perception that the road was narrower than it actually was, the mean speed of cars (but not trucks) decreased significantly; however, the effect disappeared after 30 days (Shinar, Rockwell, & Malecki, 1980).

More recently, visual illusions have been created by painting images on the road so that drivers perceive 3D speed humps (Blomberg & Cleven, 2006) and even a child chasing a ball (Staedter, 2010), but data are insufficient to warrant conclusions. The effects of these modern illusions on driving speed, accident rates, persistence (i.e., influence of images after novelty effects dissipate), and inadvertent negative effects (e.g., causing drivers to stop short, leading to rear-end collisions) will be critical measures of success.

REFERENCES

Berliner, A. (1948). *Lectures in visual psychology*. Chicago: Chicago Professional Press.

Blomberg, R. D., & Cleven, A. M. (2006). *Pilot test of Heed the Speed, a program to reduce speeds in residential neighborhoods*. Washington, DC: National Highway Traffic Safety Administration.

Boring, E. G. (1942). *Sensation and perception in the history of experimental psychology*. New York: D. Appleton-Century.

Brislin, R. W., & Keating, C. F. (1976). Cultural differences in the perception of a three-dimensional Ponzo illusion. *Journal of Cross-Cultural Psychology*, *7*, 397–412.

Bruno, N., & Franz, V. H. (2009). When is grasping affected by the Müller-Lyer illusion? A quantitative review. *Neuropsychologia*, *47*, 1421–1433.

Chadwick, P. K. (1976). Visual illusions in geology. *Nature*, *260*, 397–401.

Chapanis, A., & Mankin, D. A. (1967). The vertical–horizontal illusion in a visually rich environment. *Perception & Psychophysics*, *2*, 249–255.

Coren, S., & Girgus, J. S. (1978). *Seeing is deceiving: The psychology of visual illusions*. Hillsdale, NJ: Lawrence Erlbaum.

Day, R. H. (1965). Inappropriate constancy explanation of spatial distortions. *Nature*, *207*, 891–893.

DeLucia, P. R., & Hochberg, J. (1991). Geometrical illusions in solid objects under ordinary viewing conditions. *Perception & Psychophysics 50*, 547–554.

DeLucia, P. R., Tresilian, J. R., & Meyer, L. E. (2000). Geometrical illusions can affect time-to-contact estimation and mimed prehension. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 552–567.

Deregowski, J. B. & McGeorge, P. (2006). Oppel-Kindt illusion in three-dimensional space. *Perception*, *35*, 1307–1314.

Di Lollo, V., & Marshall, A. J. (1969). Illusion in a three-dimensional Hering figure in relation to viewing distance. *Perception & Psychophysics*, *5*, 53–55.

Fisher, G. H., & Lucas, A. (1969). Illusions in concrete situations: Introduction and demonstrations. *Ergonomics*, *12*, 11–24.

Gentaz, E., & Hatwell, Y. (2004). Geometrical haptic illusions: The role of exploration in the Müller-Lyer, vertical–horizontal, and Delboeuf illusions. *Psychonomic Bulletin & Review*, *11*, 31–40.

Georgeson, M. A., & Blakemore, C. (1973). Apparent depth and the Müller-Lyer illusion. *Perception*, *2*, 225–234.

Gibson, J. J. (1950). *The perception of the visual world*. Boston: Houghton Mifflin.

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Gillam, B. (1978). A constancy-scaling theory of the Müller-Lyer illusion. In J. P. Sutcliffe (Ed.), *Conceptual analysis and method in psychology: Essays in Honor of W. M. O'Neill* (pp. 57–70). Sydney, Australia: Sydney University Press.

Gillan, D. J., Schmidt W., & Hanowski, R. J. (1999). The effect of the Müller-Lyer illusion on map reading. *Perception & Psychophysics*, *61*, 1154–1167.

Godley, S. T., Fildes, B. N., & Triggs, T. J. (1997). Perceptual countermeasures to speed related accidents. In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics*. Vol.1: *Transportation* (pp. 387–394). Brooksfield, VT: Ashgate.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience*, *15*, 20–25.

Goodyear, W. H. (1899). Horizontal curves in Columbia University. *Architectural Record*, *9*, 82–93.

Gregory, R. L. (1963). Distortion of visual space as inappropriate constancy scaling. *Nature*, *199*, 678–680.

Heath, M., Rival, C., Neely, K., & Krigolson, O. (2006). Müller-Lyer figures influence the online reorganization of visually guided grasping movements. *Experimental Brain Research*, *169*, 473–481.

Heller, M. A., Brackett, D. D., Wilson, K., Yoneyama, K., Boyer, A., & Steffen, H. (2002). The haptic Müller-Lyer illusion in sighted and blind people. *Perception*, *31*, 1263–1274.

Julez, B. (1971). *Foundations of cyclopean perception*. Chicago: University of Chicago Press.

Leibowitz, H., Brislin, R., Perlmutrer, L., & Hennessy, R. (1969). Ponzo perspective illusion as a manifestation of space perception. *Science*, *166*, 1174–1176.

Massaro, D. W., & Anderson, N. H. (1970). A test of a perspective theory of geometrical illusions. *The American Journal of Psychology*, *83*, 567–575.

Michaels, C., & Carello, C. (1981). *Direct perception*. Englewood Cliffs, NJ: Prentice Hall.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. New York: Oxford University Press.

Müller-Lyer, F. C. (1889). Optische urteistäuschungen. *Archiv für Anatomie und Physiologie, Physiologische Abteilung, 2*(Suppl.), 263–270. (Translated in R. Day & H. Knuth [1981], The contributions of F. C. Müller-Lyer. *Perception*, *10*, 126–146.

Nijhawan, R. (1991). Three-dimensional Müller–Lyer illusion. *Perception & Psychophysics*, *49*, 333–341.

Nijhawan, R. (1995). "Reversed" illusion with three-dimensional Müller–Lyer shapes. *Perception*, *24*, 1281–1296.

Otto-de Haart, E. G., Carey, D. P., & Milne, A. B. (1999). More thoughts on perceiving and grasping the Müller–Lyer illusion. *Neuropsychologia*, *37*, 1437–1444.

Papert, S. (1961). Centrally produced geometrical illusions. *Nature*, *191*, 733.

Pirenne, M. H. (1970). *Optics, paintings, and photography*. Cambridge, UK: Cambridge University Press.

Shinar, D., Rockwell, T. H., & Malecki, J.A. (1980). The effects of changes in driver perception on rural curve negotiation, *Ergonomics*, *23*, 263–275.

Staedter, T. (2010, September 8). Optical illusion of child gets drivers to brake. *Discovery News Daily*.http://news.discovery.com/tech/optical-illusion-of-child-gets-drivers-to-brake.html

Woodworth, R. S. (1938). *Experimental psychology*. New York: Holt.

Wraga, M., Creem, S. H., & Proffitt, D. R. (2000). Perception–action dissociations of a walkable Müller–Lyer configuration. *Psychological Science*, *11*, 239–243.

Zanforlin, M. (1967). Some observations on Gregory's theory of perceptual illusions. *The Quarterly Journal of Experimental Psychology*, *19*, 193–197.

Chapter 16
Why Do Hills Look So Steep?

Frank H. Durgin and Zhi Li

INTRODUCTION

A pervasive illusion in normal human experience is the misperception of surface orientation or slant. An outdoor path that ascends a hill of some 5° is typically judged to be about 20° whether viewed from the top or bottom (Proffitt, Bhalla, Gossweiler, & Midgett, 1995). Conversely a hill that appears to be about 30° to the casual observer will typically turn out to be between 7° and 10° upon measurement. The magnitude of this error is illustrated in Figure II.16-1. With experience, skiers, hikers, engineers, and other frequent viewers of measured hills become aware of this perceptual error and may learn to make more accurate verbal estimates. However, as far as is known, the underlying perceptual bias seems to persist (Durgin & Li, 2011a). In this chapter we consider several different forms of theory that have been proposed for understanding the overestimation of geographical slant in the context of summarizing relevant findings.

To begin with ecological considerations, note that the powerful force of gravity, including both its role in surface erosion and its role in toppling leaning structures, compresses the range of ground orientations with which humans are confronted. This fact is probably quite important to understanding why surface orientation can be systematically overestimated without much cost. The statistical distribution of surface orientations in the environment may, in fact, encourage the expanded coding of surface orientation (Durgin & Li, 2011a; Proffitt, 2006). The second point is that errors in the estimation of surface orientation do not depend exclusively on visual factors.

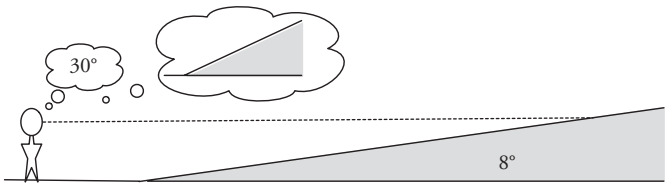


Figure II.16-1. The basic phenomenology: If a hill appears to be nearly 30°, it is probably about 8°. The perceptual error is consistent with foreshortening along the line of sight in the pictured (uphill) case, but downhill slopes also look very steep, which cannot be explained by foreshortening along the line of sight.

Estimates of underfoot ground orientation while standing on a ramp, for example, show similar patterns of overestimation—even for congenitally blind participants (Hajnal, Abdul-Malak, & Durgin, 2011). Indeed, very steep ramps (i.e., 14°) can be judged even steeper under foot (haptically) than when regarded visually (Durgin et al., 2009; Hajnal et al., 2011). Spatial bias in the haptic perception of surface slant has recently been reviewed (Durgin & Li, 2012). This chapter focuses on visual slant perception while noting similarities and differences with haptic slant perception. Several of the phenomena discussed here, such as the effect of viewing distance on slant perception, are primarily relevant to vision.

Theories of slant overestimation have fallen into the two broad categories of teleological (or *functional*) theories and mechanistic (or *incidental*) theories. Functional theories have focused on several kinds of perceptual or behavioral advantages that might arise from the exaggerated coding or representation of ground orientation. In the natural environment, sensitivity to ground orientation might be used as a basis for recognizing one's facing direction or location in a familiar geographic region (Nardi, Nitsch, & Bingman, 2010). It might also be particularly useful for taking energetics into account during route planning (Bhalla & Proffitt, 1999; Proffitt, 2006), and energy costs are highly relevant to coding the vertical dimension of space (Kammann, 1967). Perceptual error might even be used to more efficiently guide motor planning (Hajnal et al., 2011; Li & Durgin, 2012b) or simply to more efficiently represent the layout of the environment (Durgin & Li, 2011a; Durgin, Li, & Hajnal, 2010).

Incidental theories of slant misperception have included the idea that depth along the line of sight is foreshortened (Ross, 1974), which is sufficient to predict that uphill surfaces will appear too steep, that the perceived horizontal is altered in the presence of hills (O'Shea & Ross, 2007), or that perceptual biases tend to make surfaces appear more frontal than they are (Gibson, 1950). Each of these incidental theories has some measure of support, but none of them seem to fully account for the full range of observed perceptual biases. The incompleteness of these incidental theories has led some theorists to neglect their importance. But a full account of the overestimation of slant must take these facts into account as well.

This chapter argues for a hybrid theory of slant misperception that includes both functional and incidental components. Whereas the exaggeration of the vertical dimension has a clear value for layout recognition and route planning, it may be that a more general perceptual coding advantage supports both of these kinds of goals rather than privileging one or the other. The phenomenology of slant perception is relevant to this discussion.

CONSTANCY AND NONCONSTANCY IN SLANT PERCEPTION

Constancy refers to the ability of an organism to perceive one aspect of an object (such as slant) consistently despite irrelevant changes in viewing conditions (such as viewing orientation and optical distance). In this section we examine constancy and explain that there are at least two important respects in which the perception of slant is clearly not constant and three important respects in which it is surprisingly constant. In general, purely functionalist theories have limited resources for accounting for the failures of constancy described here, whereas purely incidental theories are often contraindicated by the presence of constancy.

Effects of Viewing Distance

As viewing distance from a hill surface increases, the perceived slant of the surface tends to increase (become more vertical), as illustrated in Figure II.16-2. This was demonstrated by Bridgeman and Hoover (2008) by having participants look at different fixation points along a hill of constant slope. At farther distances, perceived slants were greater. This effect is not due simply to a diminution of texture information in the distance or changes in perspective. The perceived orientations of large slanted surfaces (6°–36°) increase approximately linearly as a function of log distance even when gaze is straight ahead and monocular cues

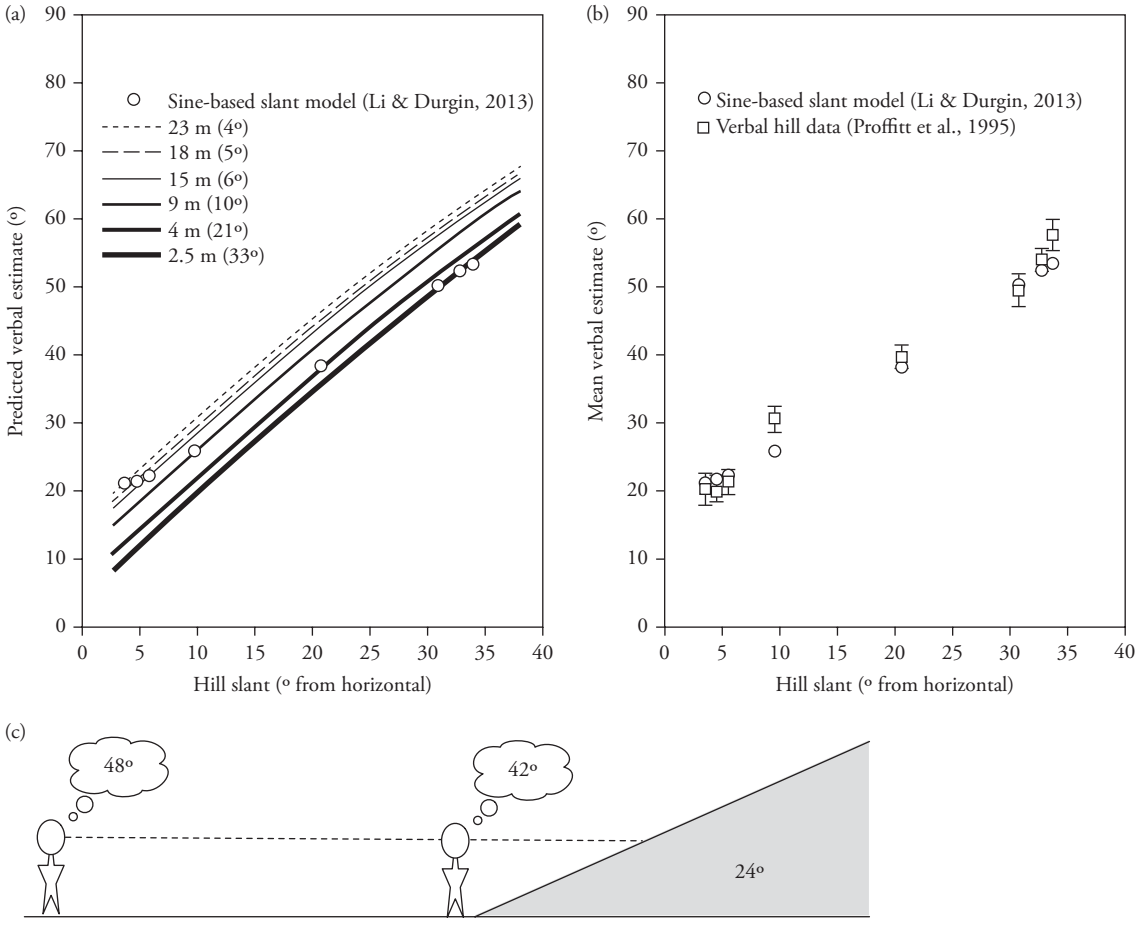


Figure II.16-2. For a farther observer, a hill will appear steeper. (a) An illustration of the empirically derived model of the effects of viewing distance on perceived slant (Li & Durgin, 2010, 2013). At each distance, slant judgments have a gain of about 1.5 (the function is primarily sinusoidal). (b) This one-parameter model fares well at predicting the hill estimation data of Proffitt et al. (1995): each white circle (in both graphs) represents the model prediction for a specific hill slant taking into account both its slant and the optical distance to hill for an observer at the base of the hill, viewing it with gaze forward. (c) The effect of distance generalizes to farther or nearer viewing: a hill of about 24° will be estimated as 42° when standing at the base (with gaze forward, the optical distance to the hill would be about 3.6 m) but as 48° when standing 10 m back from the base (i.e., 13.6 m from the hill surface at eye level).

to distance are held constant using high-resolution virtual displays (Li & Durgin, 2010). This slope nonconstancy with distance can be measured with both explicit verbal estimates of slant and with implicit slant estimates based on judgment of perceived shape on the hill surface. That is, if an L-shaped configuration of balls is simulated on a hill surface and measurement is made of the perceived ratio between frontal and sagittal arms of the L, one can use trigonometry to compute the implied surface orientation relative to the line of gaze. Such computations turn out to be roughly consistent with verbal estimates, helping to confirm that the overestimation of slant is not simply a verbal error.

In the classic hill slant estimation studies of Proffitt et al. (1995), observers were asked to make slant judgments while viewing the hills with gaze forward and standing near the base of the hill. This means that shallow hills would have been observed at a much larger optical distance than steeper hills. For example, assuming an eye height of 1.6 m, a path of 5° would only reach eye level 18 m away, whereas, a steep embankment of 33° would reach eye level within 2.5 m. To address this confound, Li and Durgin (2010) conducted their study of effects of viewing distance. When viewing distance was kept fixed, the relationship between slant and perceived slant (in the measured range of 0°–36°) turned out to be well predicted by assuming that perceived slant increases with a gain of 1.5 relative to physical slant (Li & Durgin, 2010). In fact, a gain of 1.5 is consistent with the first quarter cycle of a sinusoidal function, which also seems to approximate the function that relates physical slant to perceived slant for small surfaces in reach (Durgin & Li, 2012; Durgin, Li, et al., 2010). Li and Durgin (2013) have further developed a sine-based model as illustrated in Figure II.16-2.

Panel (a) of Figure II.16-2 illustrates the principal features of the models developed by Li and Durgin (2010, 2013): At each viewing distance, the perceived slant functions have a gain of approximately 1.5, but the effective intercepts of the functions increase with the log of viewing distance. In this case we have plotted the slant functions for viewing distances associated with each of six of the hills tested by Proffitt et al. (1995), as well as noting the specific slant shown at each distance using the sine-based model of Li and Durgin (2013, Equation 3). Panel (b) of Figure II.16-2 replots these predicted points (and three others) along with the associated means and standard errors from of Proffitt et al. Although the slant model was developed using virtual displays and implicit slant measurements (judging the aspect ratio of an L-shaped arrangement of balls on a hill surface), it provides an excellent fit to outdoor verbal estimation data.

This nonconstancy of perceived slant with respect to viewing distance is likely an incidental effect due to a failure of stereoscopic depth scaling. Although textbooks typically discuss binocular disparity as a useful depth cue only for near space, disparity information can be useful out to nearly a kilometer for surfaces that are sufficiently extended in depth (such as hills), but the scaling of stereoscopic depth information is known to show poor constancy at far distances (Palmisano et al., 2010). Many people are aware of the fact that mountains in the distance look essentially vertical. What is harder to notice (though it is observable) is that the apparent slants of hills gradually become shallower as we get closer to them.

If we consider the various functional accounts of hill misperception, it is difficult to see how any of them is strengthened by the effects of viewing distance. Perhaps a landmark could be more readily noted if its slant is exaggerated with distance, but using slant to orient to the environment would seem, on the face of it, to demand more constancy rather than less. Similarly, a functional account in terms of energetics should have trouble accounting for nonconstancies. If a climbable slope seems insurmountable when viewed in the distance, this does not seem particularly useful for route planning. The only functionalist theory that seems to be directly compatible with distance effects (i.e., not contradicted by them) is the expanded coding theory. Coding theory only need imply that slants will be exaggerated to better act upon them. Within immediate action space (several meters), the nonconstancies are both minor and fully predictable.

Failures of Constancy as Optical Slant Decreases

Li and Durgin (2009) found that downhill slants show a noticeable failure of constancy, depending on viewpoint. Using controls for viewing distance, Li and Durgin reported that downhill slants appeared steepest when the direction of view is nearly parallel to the hill surface, such as when our gaze first crests them. Many skiers have confirmed the basic observation that a downhill slope appears less steep as one gets closer to the edge. Some report that they try to maximize the sense of danger by initiating the process of launching themselves onto the hill before they have fully crested it. The apparent steepening of the hill at the point where gaze is nearly parallel with the surface may serve to signal uncertainty or risk, but it nonetheless appears to be an incidental effect rather than primarily functional because the steepening effect for downhill slants has a corresponding effect for uphill slants: When gaze is nearly parallel to slant, the visual system tends to treat gaze direction itself as the sole estimate of slant (Durgin & Li, 2011a). This effect has been observed for real and simulated downhill surfaces (Li & Durgin, 2009) as well as for small uphill simulated surfaces viewed at shallow optical slants, with gaze nearly parallel to the surface orientation (Durgin & Li, 2011a).

Evidence for Constancy With Viewing Direction

A surprising amount of constancy is evident in slant perception with respect to the lateral direction of gaze. Proffitt, Creem, and Zosh (2001) showed that, even when people looked at a hill from an oblique perspective, participants' estimates remained exaggerated. Proffitt et al. did not control for viewing distance, but their data strongly suggest that the encoding of three-dimensional slant provides substantial constancy with respect to direction of observation. For smaller, uphill surfaces, Durgin, Li, et al. (2010) reported an impressive amount of constancy as well with changes in the pitch of gaze, suggesting that coding biases in slant perception affect *geographical* slant (slant relative to gravity) rather than optical slant (slant relative to the line of sight). As shown in Figure II.16-3a, Durgin, Li, et al. had participants estimate the slants of small surfaces either with gaze forward or with gaze declined by nearly 45°. If slant misestimation were due primarily to distance foreshortening along the line of sight (also known as "frontal tendency"), then a board at about 60° from horizontal

should seem steeper with gaze forward (Fig. II.16-3a, right) than with gaze downward (Fig. II.16-3a, left), but no substantial differences were found. The function relating perceived slant to actual slant can be approximated by a sinusoid for slants within arm's reach (Li & Durgin, 2013). This effect is not verbal/numeric: the same spatial bias occurs whether numeric estimates are made relative to vertical or to horizontal. Moreover, when a forced-choice procedure was used to measure the bisection point between vertical and horizontal, surfaces of about 34° from horizontal were judged as being 45° (from horizontal and from vertical).

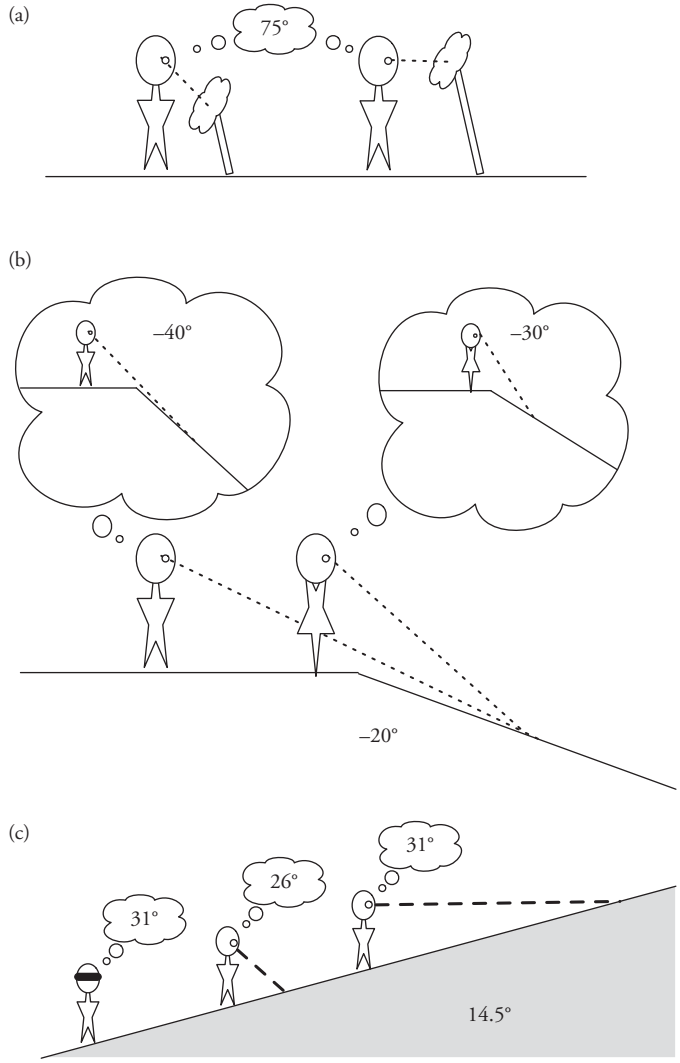


Figure II.16-3. (a) In near space, the same estimation function is found for slant whether surfaces are viewed with gaze downward or with gaze forward (Durgin, Li et al., 2010). (b) Hills look steeper when standing back from the edge so that gaze is nearly parallel to the surface of the hill. When optical slant (the angle between the center of gaze and the surface) is small, surface orientation is pulled toward the direction of gaze (Durgin & Li, 2011a; Li & Durgin, 2009). The angular declination of gaze is itself exaggerated (independent of the presence of the hill), which helps to account for the overestimation even at the edge. (c) For steep ramps, estimates given when blindfolded (haptic) more closely resemble visual estimates when gaze is forward (even when the head is bowed; Hajnal et al., 2011), than visual estimates when a nearer portion of the ramp is inspected.

Evidence for Constancy Across Surface Size

Although slant illusions are less dramatic for near surfaces at eye level, in the range of slants between 0° and about 50°, perceived orientation is expanded with a gain of about 1.5 (Durgin & Li, 2011a). This same 1.5 gain function is evident in the models of large-scale slants discussed previously. That is, the perceived orientation of small real surfaces presented under full-cue conditions are expanded in the lower half of the range by a factor of about 1.5, and exactly the same scaling factor applies to large-scale hills. Thus the perceived orientation of large-scale surfaces seems to be no different, with respect to bias, than the perceived orientation of smaller-scale surfaces. The success of the sine-based model in capturing both the hill data and the data from small surfaces shows that there is no discontinuity between small and large surfaces in slant misperception (Li & Durgin, 2013). This point, again, is consistent with the expanded coding theory because it points to a generalized coding scheme that applies to slant in general rather than exclusively to landmarks or to intended paths of travel.

Evidence for Constancy Across Modalities

As reviewed by us elsewhere, a sine-based model can apply to haptic surface perception by hand as well (Durgin & Li, 2012) and even to proprioception of hand orientation (Li & Durgin, 2012b). Thus the underlying spatial bias function in slant perception seems to be multimodal. Even in the cases that seem exceptional, further analysis suggests good calibration between modalities. For example, Durgin et al. (2009) reported that verbal estimates of a 14.5° ramp were only 26° when based on visual inspection of the ramp while standing near the base (and looking down), whereas estimates from haptic information while standing on the ramp were closer to 31°. The second (haptic) estimate corresponds well with the sine-based model estimate of visually perceived slant for a 14.5° hill when viewed with gaze forward from the base (with a viewing distance to surface of about 6.5 m). In other words, if a person were walking up such a hill and simultaneously viewing the hill with gaze forward, the visual perception of slant and the haptic perception would be aligned (both would be about 31°). In contrast, the 26° estimate can be arrived at by using the actual viewing distance to the ramp surface (about 2 m) in the model.

Summary of Slant Constancy

The evidence we have reviewed here suggests that perceived slant shows marked failures of constancy with viewing distance and with certain extreme directions of gaze with respect to surface orientation. In contrast, visually perceived slant is remarkably constant across most changes in viewing orientation, across different scales of surface size (when viewing distance is taken into account), and across different modalities. Perhaps the most significant fact about the systematic biases in slant perception is that these biases seem to be coded primarily with respect to the extrinsic reference frame specified by gravity (Durgin, Li, et al., 2010) though some incidental effects indicate that there are also consistent biases with respect to optical slant as well.

SOCIOCOGNITIVE FACTORS IN THE EVALUATION OF SLANT

A number of reports have been made suggesting that slants look steeper to people for whom they represent a greater challenge to scale. Factors that have been reported to affect the evaluation of slant include age, fitness, encumberment, fatigue (Bhalla & Proffitt, 1999), fear (Stefanucci, Proffitt, Clore, & Parekh, 2008), social support (Schnall, Harber, Stefanucci, & Proffitt, 2008), and blood sugar (Schnall, Zadra, & Proffitt, 2010). These various studies have been critiqued extensively (Durgin et al., 2009; Durgin, Hajnal, Li, Tonge, & Stigliani, 2010, 2011; Durgin, Klein, Spiegel, Strawser, & Williams, 2012; Durgin, Ruff, & Russell, 2012; Shaffer, McManama, Swank, & Durgin, 2013). In some cases, such as a study of the elderly, the originally published data actually contradicted the hypothesis: The elderly gave lower estimates for most hills, but this was not made evident in the initial report. In other cases, such as a study of fitness, confounding factors (e.g., sex differences in slant estimation) were not taken into account in the analyses (see Durgin, Hajnal, et al., 2010 for a discussion). In yet others concerned with fear, subjects may have been excluded from analysis in a manner that inadvertently biased the results (see Durgin et al., 2009, for a discussion).

The critiques of the studies of encumberment are worth reviewing briefly here. If participants in a study are simply asked to wear a heavy backpack, they tend to give higher estimates for hills than do nonencumbered participants. However, they also tend to report that they thought the backpack was supposed to make the hill look steeper (Durgin et al., 2009; Durgin, Klein, et al., 2012). If the backpack is instead presented as carrying equipment essential to the conduct of the experiment (Durgin et al., 2009; Durgin, Ruff, et al., 2012; Shaffer et al., 2013), or participants are told not to let themselves be influenced by the backpack (Durgin, Klein, et al., 2012), the judgments that they give tend to be identical to those of nonencumbered participants. In general, the use of heavy backpacks that are transparently intended to increase estimates of slant seem to produce social pressure on participants to elevate their slant estimates. Susceptibility to that social pressure can be mitigated by social support (Schnall et al., 2008) and may be exacerbated by low blood sugar (Durgin, Klein, et al., 2012; Schnall et al., 2010; Shaffer et al., 2013). Durgin, Klein, et al. (2012) found that the effects of blood sugar disappeared when participants were simply told to ignore the heavy backpack they were required to wear during such studies. Shaffer et al. (2013) showed that low blood sugar produces opposite effects (lower estimates) if participants believe the drink (which they assume contained sugar) was supposed to make the hill look shallower. Disputes over these sorts of controversial findings remain lively in the literature (e.g., Firestone, 2013). It is very clear that social compliance can have powerful effects on judgments and should be taken into consideration in studies that are ostensibly of perception.

The idea that slant overestimation provides a means for the direct perception of the energetic affordances of the environment (Proffitt, 2006) is quite a clever one. However, much of the evidence amassed in support of this theory has proven problematic, and no direct connection seems to exist between energetics and slant perception (e.g., Shaffer & Flint, 2011). A surviving tenet of this view, however, is that geometric accuracy might not be the proper goal of perceptual representation. This is also a tenet of expanded scaling theory, and if the exaggeration of perceived slant incidentally helps people to more reliably evaluate the energetic affordances of the environment for the purpose of route planning, so much the better.

THE ROLE OF THE PERCEIVED DIRECTION OF GAZE

Because slant is defined relative to a gravitational reference frame (i.e., the horizontal plane and the vertical vector of gravity that is normal to horizontal), errors in perceived slant could come about if the presence of a hill produced a distortion in the perception of the horizontal plane. O'Shea and Ross (2009) have provided evidence for such effects in the presence of large-scale mountains (see also Matin & Li, 1992), and Ooi and He (2007) have suggested that the ground plane itself is perceived as being tilted upward. The magnitude of such effects, however (about 3°–5°), is insufficient to account for the very large magnitudes of distortion in perceived slant.

Recently a much more dramatic distortion in perceived gaze direction has been documented that seems more consistent with the overestimation of slant. Specifically, much as the perceptual gain for perceived slants (less than 45°) is about 1.5 when distance is fixed, Li and Durgin (2009; Durgin & Li, 2011a) have used a variety of methods to document that the perceived declination of gaze is also coded with a gain of about 1.5. Durgin and Li (2011a; Li & Durgin, 2012a) have proposed that angular variables relevant to the pitch axis are coded on an expanded scale so as to increase their precision because they are highly relevant for action.

Gaze declination relative to the horizon is a powerful cue to distance (Messing & Durgin, 2005; Ooi, Wu & He; 2001; Sedgwick, 1986; Wallach & O'Leary, 1982), because on level ground (and most of the spaces we deal with are within 5° of level) it provides highly reliable proprioceptive cue to ground distance. Given the bandwidth limitations of neural transmission, coding this angular variable on an expanded scale would preserve greater precision relevant for the control of action. Thus a functional account of slant overestimation derives in part from the idea that angular distortions are present in perceptual experience in order to maintain precision for action. The expanded scaling of perceived gaze declination has been measured implicitly (Durgin & Li, 2011a; Li & Durgin, 2009) based on slant estimates viewed along different lines of sight; it has been measured directly with balls suspended in the air or placed along the ground, and it has been measured by means of a bisection task (Durgin & Li, 2011a; see also Durgin & Li, 2011b). In all cases the gain was found to be approximately 1.5. Studies of verbal distance estimation among nonexperts tend to suggest linear compression of perceived ground distance by a factor of 0.7 to 0.8 (Loomis & Philbeck, 2008), consistent with a misperception of gaze direction, as illustrated in Figure II.16-4a. Rather than hills looking steep because of distance foreshortening, it may be that ground distances are underestimated because crucial

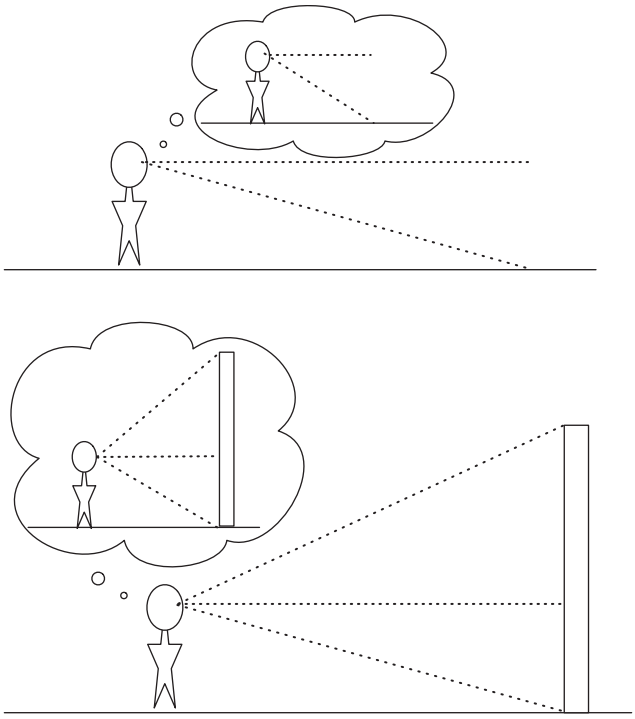


Figure II.16-4. (a) Independent of the presence of a ground surface, perceived gaze declination (within 45° of horizontal) is exaggerated by a gain of 1.5 (Durgin & Li, 2011a; Li & Durgin, 2009). (b) Asked to set themselves the same distance from a pole as the pole is high, participants set themselves much too far away (Higashiyama & Ueyama, 1988). The mismatch between their perceptual experience and the physical situation can be predicted by the scaling of the perceived angular declination of gaze (Li et al., 2011).

angular variables, such as the angle of gaze declination, are systematically misperceived.

Strikingly, this expanded scaling of perceived angular declination of gaze (along with a concomitant scaling of optical slant) can predict not only downhill slant perception (Li & Durgin, 2009; see Fig. II.16-3, bottom) but also systematic errors in the comparison of distance and height. For example, Higashiyama and Ueyama (1988) developed a task requiring participants to place themselves at the same distance from objects (such as poles) as those objects were high. Participants placed themselves much too far away, as illustrated in Figure II.16-4b. Li, Phillips, and Durgin (2011) recently replicated this experiment and extended it to show that the exact pattern of results reported by Higashiyama and Ueyama could be predicted by a parameter-free geometric model in which the previously measured angular declination gain of 1.5 was assumed. This suggests that hill misperception is part of a larger pattern of angular distortions that affect the perception of surface layout generally.

THEN WHY DO HILLS LOOK SO STEEP?

Slant misperception is dramatic. An editor at a journal once challenged the statement that Lombard Street in San Francisco is on a hill that is only 15° in slope. He said he had checked the Internet and found that the true value is 31°. He was correct that many sites on the Internet report a value of 31°. In fact, the tangent of 15° is 0.31, and so the grade of the hill is 31% (a 100% grade would be a 45° slope). Since 15° is simply unbelievable for anyone who has been to Lombard Street, many websites simply report this as 31° (which still seems too low compared to the perceptual phenomenology). The hill is so steep that Lombard Street winds back and forth across it so as to reduce the effective slant of the road to 10°—still quite steep when walking up it!

As Marr (1982) pointed out, there are many different forms of answers to questions about explanation, including the functional and the mechanistic. This chapter has considered the phenomenology of hill perception while discussing a variety of theories that have been proposed. One answer to the question in the title might remain the one put forth by Kammann (1967): because of gravity. The environment in which we have evolved is laid out such that vertical extents and slants are relatively tiny compared to horizontal extents. Expanding the vertical scaling of such an environment might produce many cognitive advantages, even if that expansion is done in angular terms. A second, mechanistic answer remains: because of a loss of reliable information for depth along the line of sight. This latter answer addresses the failure of slant constancy with changes in viewing distance, but it cannot be the whole story, because it does not account for the misperception of downhill slant or the relative constancy of perceived slant with large changes in angle of regard. This chapter has emphasized that the misperception of hills is probably part of a larger family of biases in the perception of angular variables that includes the misperception of small surfaces in reach and may even include the misperception of ground distance and height based on multiplicative biases in perceived angular deviations of visual direction from horizontal.

Although this review has taken sides on some controversies, there is not room to address controversies over the interpretation of different measurement techniques for evaluating perceived slant (e.g., Coleman & Durgin, 2014; Creem & Proffitt, 1998; Durgin, 2013; Durgin, Hajnal, et al., 2010; Durgin, Li, et al., 2010; Li & Durgin, 2010, 2011, 2013; Shaffer, McManama, Swank, Williams, & Durgin, 2014; Stigliani, Li, & Durgin, 2013; Taylor-Covill & Eves, 2013; Witt & Proffitt, 2007). These controversies focus around the question of whether using a haptic matching task to measure slant provides a route to a separate (undistorted) dorsal stream representation.

It is therefore worth making one final point about the nonconsequences of slant misperception. Our actions seem to be coded in the same perceptual space as everything else (Powers, 1973). This means that acting with accuracy in a perceptually distorted world requires only that actions be calibrated to the same distortions (Durgin, 2009). For example, because proprioception of hand orientation is distorted with precisely the same function as the haptic and visual perception of surfaces (Li & Durgin, 2012b), the perceptual distortions documented here are transparent to our action systems. We can live and act effectively in a distorted visual world. Because the distortion is fairly stable, and the correlations between motor signals and sensory signals are maintained, even the effects of nonconstancy with distance can be predicted and therefore ignored in our normal perceptual experience. An important part of the

answer to why hills (and even small surfaces in reach) look so steep is therefore: Why not?

ACKNOWLEDGMENTS

Preparation of this chapter was supported by Award Number R15 EY021026 from the National Eye Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Eye Institute or the National Institutes of Health.

REFERENCES

Bhalla, M., & Proffitt, D. R. (1999). Visual-motor recalibration in geographical slant perception. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(4), 1076–1096.

Coleman, A., & Durgin, F. H. (2014). Egocentric reference frame bias in the palmar haptic perception of surface orientation. *Psychonomic Bulletin & Review*, *21*(4), 955–960.

Creem, S. H., & Proffitt, D. R. (1998). Two memories for geographical slant: Separation and interdependence of action and awareness. *Psychonomic Bulletin & Review*, *5*(1), 22–36.

Durgin, F. H. (2009). When walking makes perception better. *Current Directions in Psychological Science, 18*(1), 43–47.

Durgin, F. H. (2013). What do hands know about hills? Interpreting Taylor-Covill and Eves (2013) in context. *Acta Psychologica, 144*(2), 451–458.

Durgin, F. H., Baird, J. A., Greenburg, M., Russell, R., Shaughnessy, K., & Waymouth, S. (2009). Who is being deceived? The experimental demands of wearing a backpack. *Psychonomic Bulletin & Review, 16*(5), 964–969.

Durgin, F. H., Hajnal, A., Li, Z., Tonge, N., & Stigliani, A. (2010). Palm boards are not action measures: An alternative to the two-systems theory of geographical slant perception. *Acta Psychologica, 134*(2), 182–197.

Durgin, F. H., Hajnal, A., Li, Z., Tonge, N., & Stigliani, A. (2011). An imputed dissociation might be an artifact: Further evidence for the generalizability of the observations of Durgin et al. 2010. *Acta Psychologica, 138*(2), 281–284.

Durgin, F. H., Klein, B., Spiegel, A., Strawser, C. J., & Williams, M. (2012). The social psychology of perception experiments: Hills, backpacks, glucose and the problem of generalizability. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(6), 1582–1595.

Durgin, F. H., & Li, Z. (2011a). Perceptual scale expansion: An efficient angular coding strategy for locomotor space. *Attention, Perception & Psychophysics, 73*(6), 1856–1870.

Durgin, F. H., & Li, Z. (2011b). The perception of 2D orientation is categorically biased. *Journal of Vision*, *11*(8), 13.

Durgin, F. H., & Li, Z. (2012). Spatial biases and the haptic experience of surface orientation. In A. El Saddik (Ed.), *Haptics rendering and applications* (pp. 75–94). Maastricht, The Netherlands: INTECH.

Durgin, F. H., Li, Z., & Hajnal, A. (2010). Slant perception in near space is categorically biased: Evidence for a vertical tendency. *Attention, Perception & Psychophysics, 72*(7), 1875–1889.

Durgin, F. H., Ruff, A. J., & Russell, R. (2012). Constant enough: On the kinds of perceptual constancy worth having. In G. Hatfield & S. Allred (Eds.), *Visual experience: Sensation, cognition and constancy* (pp. 87–102). Oxford: Oxford University Press.

Firestone, C. (2013). How "paternalistic" is spatial perception? Why wearing a heavy backpack doesn't—and couldn't—make hills look steeper. *Perspectives on Psychological Science, 8*(4), 455–473.

Gibson, J. J. (1950). The perception of visual surfaces. *American Journal of Psychology, 63*, 367–384.

Hajnal, A., Abdul-Malak, D. T., & Durgin, F. H. (2011). The perceptual experience of slope by foot and by finger. *Journal of Experimental Psychology: Human Perception and Performance, 37*(3), 709–719.

Higashiyama, A., & Ueyama, E. (1988). The perception of vertical and horizontal distances in outdoor settings. *Perception & Psychophysics, 44*(2), 151–156.

Kammann, R. (1967). Overestimation of vertical distance and slope and its role in the moon illusion. *Perception & Psychophysics, 2*(12), 585–589.

Li, Z., & Durgin, F. H. (2009). Downhill slopes look shallower from the edge. *Journal of Vision, 9*(11), 6.

Li, Z., & Durgin, F. H. (2010). Perceived slant of binocularly viewed large-scale surfaces: A common model from explicit and implicit measures. *Journal of Vision*, *10*(14), 13.

Li, Z., & Durgin, F. H. (2011). Design, data and theory regarding a digital hand inclinometer: A portable device for studying slant perception. *Behavior Research Methods, 43*(2), 363–371.

Li, Z., & Durgin, F. H. (2012a). A comparison of two theories of perceived distance on the ground plane: The angular expansion hypothesis and the intrinsic bias hypothesis. *i-Perception*, *3*(5), 368–383.

Li, Z., & Durgin, F. H. (2012b). Manual matching of perceived surface orientation is affected by arm posture: Evidence of calibration between proprioception and visual experience in near space. *Experimental Brain Research, 216*(2), 299–309.

Li, Z., & Durgin, F. H. (2013). Depth compression based on misscaling of binocular disparity may contribute to angular expansion in perceived optical slant. *Journal of Vision 13*(12), 3.

Li, Z., Phillips, J., & Durgin, F. H. (2011). The underestimation of egocentric distance: Evidence from frontal matching tasks. *Attention, Perception & Psychophysics*, *73*(7), 2205–2217.

Loomis, J. M., & Philbeck, J. W. (2008). Measuring spatial perception with spatial updating and action. In R. L. Klatzsky, B. MacWhinney, & M. Behrmann (Eds.), *Embodiment, ego-space, and action* (pp 1–43). New York: Taylor & Francis.

Marr, D. (1982). *Vision*. New York: Freeman.

Matin, L., & Li, W. (1992). Visually perceived eye level: Changes induced by a pitched-from-vertical 2-line visual field. *Journal of Experimental Psychology: Human Perception and Performance, 18*(1), 257–289.

Messing, R. M., & Durgin, F. H. (2005). Distance perception and the visual horizon in head-mounted displays. *ACM: Transactions on Applied Perception*, *2*, 234–250.

Nardi, D., Nitsch, K. P., & Bingman, V. P. (2010). Slope-driven goal location behavior in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 36*(4), 430–442.

Ooi, T. L., & He, Z. J. (2007). A distance judgment function based on space perception mechanisms: Revisiting Gilinsky's (1951) equation. *Psychological Review*, *114*(2), 441–454.

Ooi, T. L., Wu, B., & He, Z. J. (2001). Distance determined by the angular declination below the horizon. *Nature*, *414*(6860), 197–200.

O'Shea, R. P., & Ross, H. E. (2007). Judgments of visually perceived eye level (VPEL) in outdoor scenes: Effects of slope and height. *Perception, 36*(8), 1168–1178.

Palmisano, S., Gillam, B., Govan, D. G., Allison, R. S., & Harris, J. M. (2010). Stereoscopic perception of real depths at large distances. *Journal of Vision*, *10*(6), 19.

Powers, W. T. (1973). *Behavior: The control of perception.* Chicago: Aldine.

Proffitt, D. R. (2006). Embodied perception and the economy of action. *Perspectives on Psychological Science, 1*(2), 110–122.

Proffitt, D. R., Bhalla, M., Gossweiler, R., & Midgett, J. (1995). Perceiving geographical slant. *Psychonomic Bulletin & Review, 2*(4), 409–428.

Proffitt, D. R., Creem, S., & Zosh, W. (2001). Seeing mountains in molehills: Geographical slant perception. *Psychological Science, 12*(5), 418–423.

Ross, H. E. (1974). *Behavior and perception in strange environments* London: Allen & Unwin.

Schnall, S., Harber, K. D., Stefanucci, J., & Proffitt, D. R. (2008). Social support and the perception of geographical slant. *Journal of Experimental Social Psychology, 44*(5), 1246–1255.

Schnall, S., Zadra, J. R., & Proffitt, D. R. (2010). Direct evidence for the economy of action: Glucose and the perception of geographical slant. *Perception, 39*(4), 464−482.

Sedgwick, H. A. (1986). Space perception. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.). *Handbook of perception and human performance* (pp. 21.1–21.57). New York: Wiley.

Shaffer, D. M., & Flint, M. (2011). Escalating slant: Increasing physiological potential does not reduce slant overestimates. *Psychological Science*, *22*(2), 209–211.

Shaffer, D. M., McManama, E., Swank, C., & Durgin, F. H. (2013). Sugar and space? Not the case: Effects of low blood glucose on slant estimation are mediated by beliefs. *i-Perception, 4*(3), 147–155.

Shaffer, D. M., McManama, E., Swank, C., Williams, M., & Durgin, F. H. (2014). Anchoring in action: Manual estimates of slant are powerfully biased toward initial hand orientation and are correlated with verbal report. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1203–1212.

Stefanucci, J. K., Proffitt, D. R., Clore, G. L., & Parekh, N. (2008). Skating down a steeper slope: Fear influences the perception of geographical slant. *Perception, 37*(2), 321–323.

Stigliani, A., Li, Z., & Durgin, F. H. (2013). Humans have precise knowledge of familiar geographical slants. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(6), 1966–1973.

Taylor-Covill, G. A. H., & Eves, F. F. (2013). The accuracy of "haptically" measured geographical slant perception. *Acta Psychologica, 144*(2), 444–450.

Wallach, H., & O'Leary, A. (1982). Slope of regard as a distance cue. *Perception & Psychophysics, 31*(2), 145–148.

Witt, J. K., & Proffitt, D. R. (2007). Perceived slant: A dissociation between perception and action. *Perception*, *36*, 248–257.

Chapter 17
"Shape From Smear"

An Illusion of 3D Shape, Made by Finger-Painting With Noise

Roland W. Fleming and Daniel Holtmann-Rice

INTRODUCTION

In everyday language, the word "illusion" usually implies some kind of fantastical hallucination or erroneous sensory experience, like seeing a lake in the middle of a desert or perceiving two lines to be different lengths when in fact they are the same. This compendium is packed full of mind-bending sensory paradoxes that truly match this everyday concept of what an illusion is. However, in this chapter, by comparison, the "illusion" we present is utterly mundane. In fact, the experience is so obvious that readers would be forgiven for not wanting to call it an illusion at all—and in a sense they would be right, because we believe it results from normal neural processes that allow our brain to reconstruct the three-dimensional (3D) shape of objects in our environment. Nevertheless, the fact that we see it the way we do can give us key insights into visual processing, as described in our article using such images (Fleming, Holtmann-Rice and Bülthoff, 2011).

A simple example of the "illusion" is shown in Figure II.17-1. Most people see the pattern as a smoothly curved, irregular 3D object, whose surface is coated with some kind of random patterns. The impression of a 3D shape is so clear and obvious that there does not seem to be anything particularly surprising about it. It just looks like some sort of computer graphics rendering (and not very good computer graphics at that). But this was not created like a normal rendering, and, in fact, to understand why we call this an illusion at all, one must understand how we made the image.

TEXTURE COMPRESSION AND THE MAKING OF SHAPE FROM SMEAR

The image was created by starting with a square of random noise, much like the "snowstorm" patterns one used to see on a television set when the antenna was unplugged. Every point in the noise is assigned a random intensity value, so it contains no discernable structure, and certainly nothing that that looks like a 3D surface. An example is shown in Figure II.17-2. It does not look like much of anything in particular.

To understand what we did next, it is important to understand how texture markings on surfaces behave when projected onto the retina, as it is this process that makes the pattern in Figure II.17-1 appear 3D. Many objects in our natural environment have patterns or texture markings on their surfaces. Examples of common textures include wood grain, leopard spots, or the checks on a shirt. What defines these as textures is the presence of some kind of statistical process in which basic texture elements are spread out or repeated across the surface. Not all the elements have to be identical, and they do not have to be arranged regularly (imagine, for example, the distinctive "texture" created by daisies and blades of grass on a garden lawn). However, in order for a visual pattern to count as a texture, there needs to be some sense in which different samples (i.e., different patches) of the texture contain the same kinds of elements arranged according to the same kind of process.

By and large, most texture patterns are roughly evenly spaced across an object's surface. For example, the small cracks in the surface of a piece of leather are roughly the same size and divide the leather up into small cells that are roughly the same size, wherever one looks on the surface. However, when we view a 3D object, different points on the surface lie at different distances and angles relative to the line of sight, and this affects the appearance of the texture when it is projected onto the retina. The process of projection causes the texture to appear compressed or distorted in various ways, and this provides important information about 3D shape. We briefly describe how this works and then return to how we made the image in Figure II.17-1.

There are two main ways in which the texture is affected by projection. The first is an effect of the distance of the surface from the observer's viewpoint. When a textured surface is moved further away from the eye, the patterns appear smaller in the image; when it is moved closer, the patterns appear larger. This is demonstrated in the top row of Figure II.17-3. This is a simple consequence of perspective projection—just as objects on the horizon appear much smaller than those that are close by.

The fact that texture changes size depending on the distance of the surface means that the brain could use the

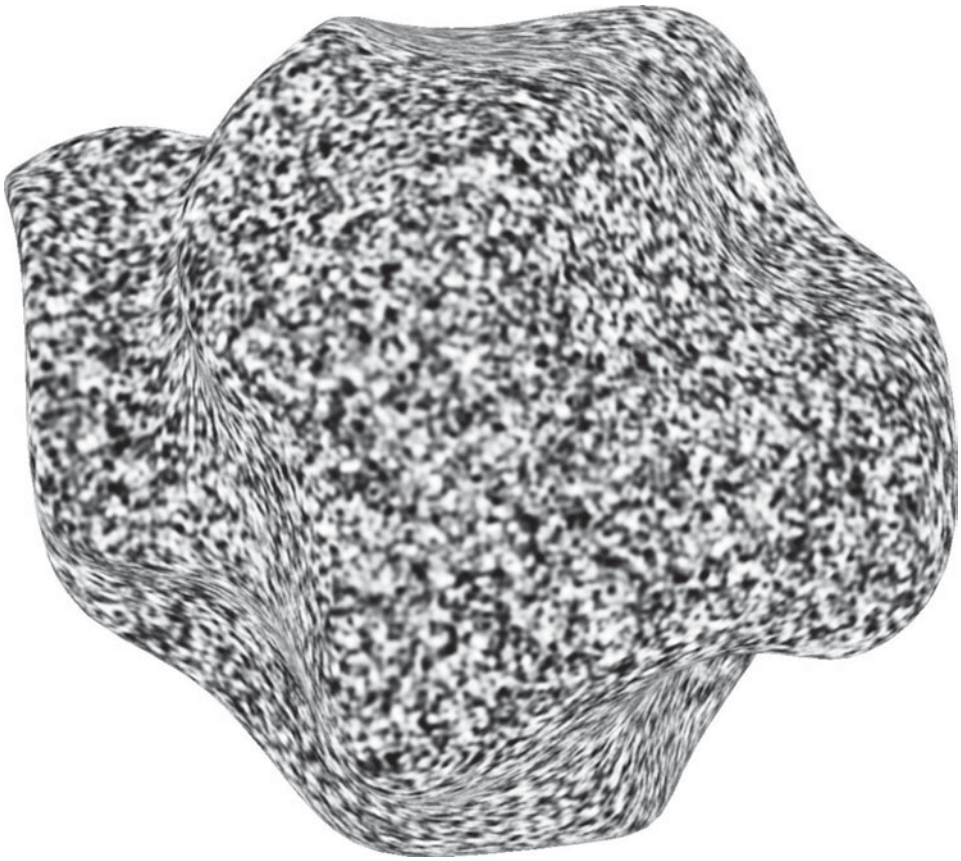


Figure II.17-1. An example of the illusion we call "shape from smear." You probably see a smoothly curved 3D shape, even though the image is just made out of 2D patterns. That is a perfectly normal consequence of the way our brains estimate 3D shape from texture patterns. Images: Roland W Fleming and Daniel Holtmann-Rice.

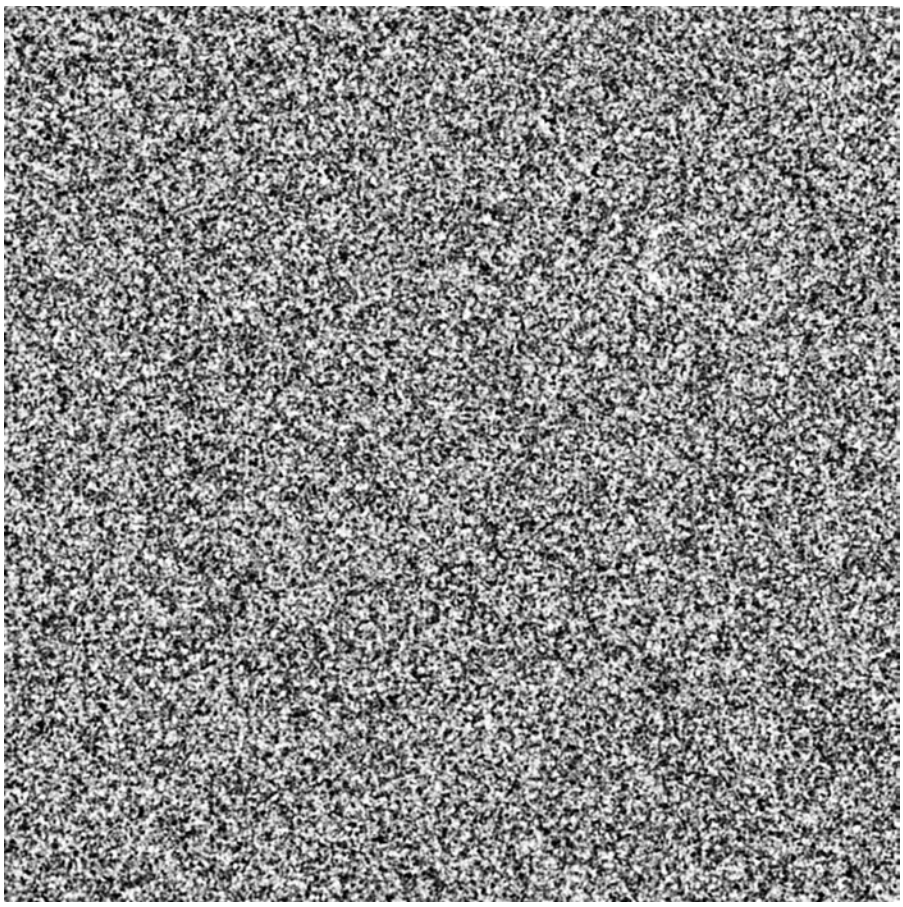


Figure II.17-2. The image in Figure II.17-1 started life as a pattern of random noise, like this. Noise does not look very much like a 3D object. Images: Roland W Fleming and Daniel Holtmann-Rice.

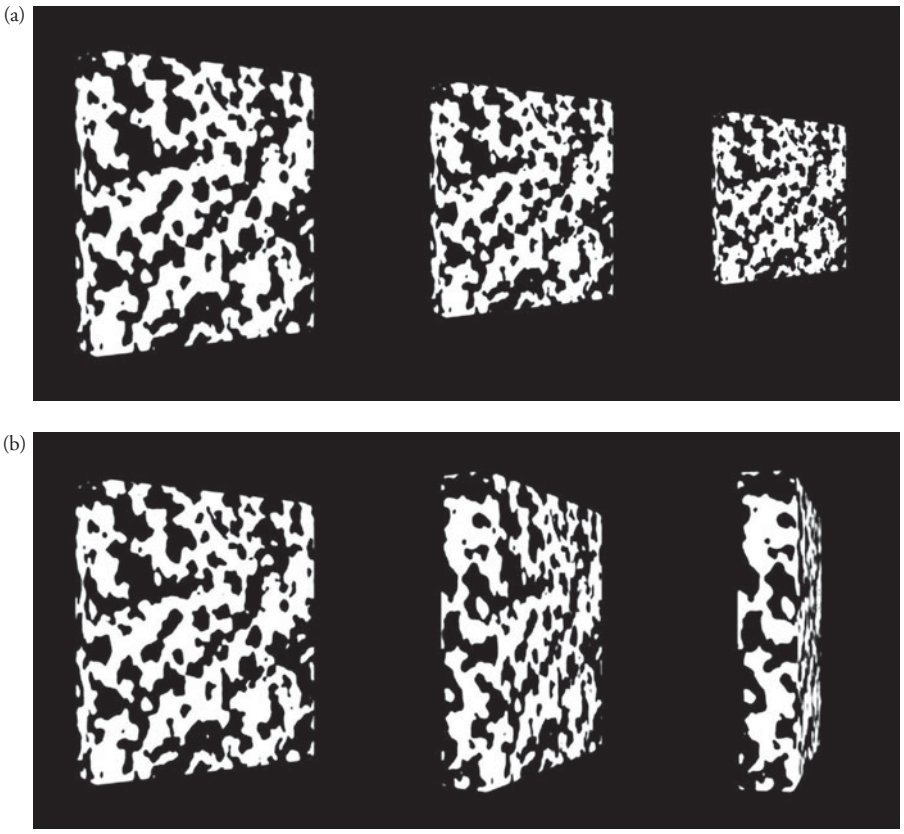


Figure II.17-3. How texture is compressed when it is projected into the retina. The top row shows the effects of distance on the projected texture. The texture gets smaller the further the surface is from the viewer. The bottom row shows the effects of slant on the projected texture. The more slanted the surface is relative to the line of sight, the more compressed the texture appears in the direction of the slant. Images: Roland W Fleming and Daniel Holtmann-Rice.

relative size of the texture patterns to work out the relative distance of surfaces. Put another way, texture compression provides a "visual cue" to depth. Under most circumstances, however, this cue is not very powerful, because the variations in depth within an object are usually much smaller than the distance from the eye to the object as a whole. Indeed, this cue disappears completely under orthographic projection (i.e., when the projected rays are parallel, like an extreme telephoto lens), and yet textured objects still appear 3D, suggesting the brain does not rely heavily on this cue.

The second effect of projecting texture into the image provides a much more powerful and important source of information about 3D shape, which depends on the angle of the surface relative to the line of sight. When a textured surface is viewed from head on, the texture appears evenly spaced. But when the surface is viewed from a slanted angle, the texture patterns appear bunched up or compressed along the direction of the slant. This can be seen clearly in the bottom row of Figure II.17-3, where the slant of the surface progressively increases from left to right.

Importantly, whereas distance compresses the texture uniformly in all directions, surface slant compresses texture only in the direction of the slant. Note that when the surface is highly slanted (right-hand side of the figure), the texture gets bunched up only in the horizontal direction, so that the texture appears in the image as an aggregation of vertically oriented streaks.

The fact that this second type of texture compression is directional means that it provides two key sources of information about the orientation of the surface. First, the degree of compression tells the brain how slanted the surface is. Second, the orientation of the compression tells the brain the direction in which the surface is slanted. This provides a very useful source of information about 3D object shape, especially near the edges of objects, where the front face of the object curves out of sight as it wraps around to the back face (where it is no longer visible). At the locations where the object curves out of sight, the surface becomes very slanted relative to the observer's line of view, so the degree of compression is large, irrespective of the size or distance of the object. This is one of several reasons why compression resulting from the surface slant is a more powerful shape cue.

For planar surfaces (Fig. II.17-3), changes in the amount of compression are only caused by the changing distance of points on the plane from the viewer, since the slant of the plane is constant. However, for more complex curved objects (Fig. II.17-4), as we move from point to point across the surface, the slant varies continuously in both direction and magnitude. Locations that face the viewer are compressed *isotropically* (i.e., the same amount in all directions), whereas locations that are highly slanted lead to *anisotropic* texture compression (i.e., compression along a specific direction).

In Figure II.17-4 we map out the amount of compression caused by both the distance and slant of the surface at each location on the surface. In the "depth map," surface points that are distant from the observer are colored dark grey, whereas points that are close to the observer are colored light grey. In the "slant map" image, the colors

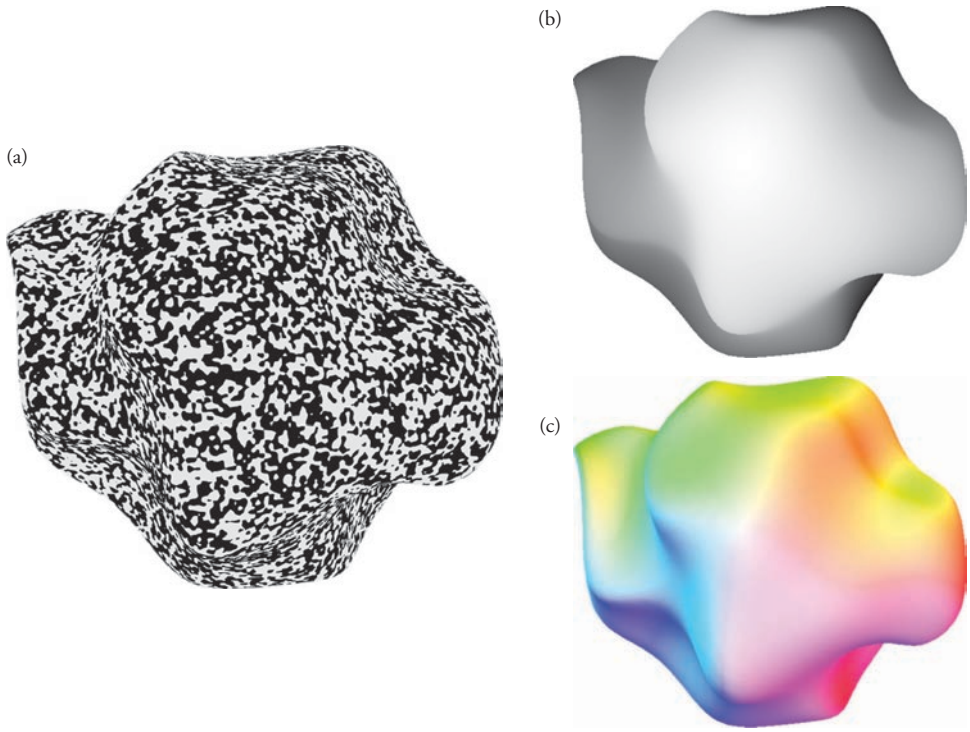


Figure II.17-4. Texture compression for an irregular 3D shape. The image on the left was created by ray tracing an object coated in random texture. The two maps on the right indicate the surface orientations and depths of the surface, which together determine the amount by which the texture appears compressed at each location in the image. For the "depth map," darker locations indicate more distant surface points and brighter locations indicate nearer surface points. For the "slant map," hue indicates the direction of slant and color saturation indicates the degree of slant, with white signifying frontoparallel surface locations. Images: Roland W Fleming and Daniel Holtmann-Rice.

represent both the direction and degree of slant of the surface. Specifically, the *hue* of the color (i.e., its shade, such as blue, red, or green) represents the direction of the slant, and the *saturation* of color represents the degree of slant, with vibrant colors indicating that the surface is highly slanted and washed-out pastel tones indicating the surface is close to facing toward the viewer.

These two maps describe exactly how the texture is compressed in the image. In fact, importantly, the maps describe how *any* texture would be compressed in the image, as long as the texture is homogenous in scale and random in all directions (e.g., leaves, pebbles, or granite), rather than directional (e.g., zebra stripes or wood grain). This is important as it means that we have abstracted something about the way the texture is affected by the 3D shape of the surface, irrespective of what the texture is. In other words, whether the texture consists of big leaves or small pebbles, perfectly round coins or irregularly shaped grains of soil, the texture will be compressed in exactly the same way: by the amounts and in the directions indicated by the depth and slant maps. This is the key information that the brain can use for estimating 3D shape, and this is the information that makes the object in Figure II.17-4 appear 3D.

Before we get back to explaining how we made the image in Figure II.17-1—the one we called an illusion—it is worth briefly discussing the 3D appearance of the texture object in Figure II.17-4. The image of the object in Figure II.17-4 was created using computer graphics. The surface of a 3D object was coated evenly with texture, and then a ray-tracing algorithm was used to render the image from a specific viewpoint. This resulting image is as close as computationally possible to accurately simulating the physical behavior of texture when it is projected into the retina. The resulting image looks clearly 3D, even though the image itself is, of course, flat. However, most researchers working on visual perception would not call this an illusion—otherwise all of vision would be classified as an illusion, because all of vision is based on two-dimensional (2D) retinal images. The image looks 3D because it contains exactly the information that the brain uses in everyday life to work out the 3D shape of objects from texture compression. Thus, just as we do not normally call a photograph "an illusion," this image would not normally be considered an illusion.

By contrast, the image in Figure II.17-1 was not created by accurately simulating the physics of projection. Instead, the image was created using entirely 2D image-processing tricks, a bit like the filters and effects found in photo-editing software. The way we did this is as follows.

First, we chose a 3D object that we wanted to simulate. Viewing the shape from a chosen viewpoint, we calculated the outline and the slant map for that shape. The next step was to take a random noise pattern and crop out only the portion of the noise that was inside the outline of the object. On its own, this did little to produce an impression of three-dimensionality. Clearly we needed to do something more to the noise if we wanted to make it look 3D.

Simulating the effects of slant requires two steps. Recall that in slanted regions of the shape, compression shrinks elements of the texture in the direction of slant. To mimic this effect of slant on the size of the texture elements, we blurred the noise pattern in areas where the amount of slant was low (where the texture elements should be big). This blurring pooled the fine dots in the noise, turning them into larger light and dark spots. In areas where the amount of slant was higher, less blurring was applied, leaving more of the fine structure of the noise and making the pattern elements look smaller. While this process produced an image conveying information about the amount of slant

at each point, the directional component of the compression was not included.

The next step was the critical one for making the noise appear 3D. We used two 2D image-processing operations called "line integral convolution" and "anisotropic diffusion" to change the properties of the texture at each location in the image. These processes "smear" the underlying texture along a given direction at each point in the image. It is a bit like finger-painting. Imagine that the dots in the noise are wet paint. Moving a finger along a given direction spreads the paint out along that direction, turning the little dots into extended streaks. Because we can do this in different directions at each point in the image, we can create a pattern in which the dots are smeared out into extended lines and curves that run smoothly from location to location across the image. This is what creates the impression of a 3D shape.

To be more specific, we smeared the noise along the directions perpendicular to the directions in the slant map. This was intended to mimic the directional effects of slant on texture compression, although, importantly, this is not the same as a physically accurate rendition, as we discuss later.

The outcome of these image-processing tricks is the image shown in Figure II.17-1. Clearly, blurring and smearing flat noise in these specific ways created a strong illusion of 3D shape. Even though the image processing was entirely 2D in nature, the resulting image appears vividly 3D.

How is this different from rendering? The crucial difference is that in a rendering the process of projection can only ever *compress* the texture in the image along the direction of the slant. For a real surface, no matter how slanted it is, the texture elements can never appear longer in the image than they are in real life. By contrast, here we did not compress the texture at all. Instead we have smeared it out along the orthogonal directions in the image. This creates parallel streaks that are not physically accurate renditions of texture markings from the surface into the image, because they are elongated in one direction rather than compressed in the opposite direction. Despite this, the image still appears 3D. In other words, even though the image is not an accurate rendition of a textured surface, it nevertheless produces the subjective impression of a specific 3D shape. This is why we call it an illusion. And, like many other illusions, the fact that it works can give us some useful insights into the way the brain processes visual information to reconstruct the surrounding world.

HOW WE THINK THE ILLUSION WORKS

The fact that Figure II.17-1 looks 3D indicates that somehow we have stimulated the appropriate machinery in the brain. But what machinery? Although this illusion does not allow us to work out the entire brain pathway involved in reconstructing shape from texture, it does suggest a role of a specific class of cells that are known to be present in primary visual cortex. In the 1950s and 1960s, Hubel and Wiesel conducted a seminal series of physiological experiments on the behavior of cells in cat visual cortex, which eventually won them the Nobel Prize (see, for example, Hubel and Wiesel, 1959). They discovered that many cells respond to bright–dark boundaries (i.e., edges or lines in the image) at specific locations and orientations on the retina. If the edge or line is presented outside a cell's receptive field on the retina, the cell does not respond (this is what defines a cell's receptive field). Importantly, if the edge or line is inside the receptive field but has the incorrect orientation, the cell also does not respond—it only responds to a narrow range of orientations. In other words, the cell is "tuned" to specific orientations of edges within its receptive field. For each location in the image, there are many cells with different orientation preferences, which respond to different edges presented at that location. This way, the brain encodes all possible orientations at every location in the retina.

Because these cells respond to edges and bars, for a long time they were referred to as "edge-detectors" and were thought to detect the edges or boundaries of objects in the image. However, the cells respond to *any* oriented bright–dark pattern in the image that lines up with their preferred orientation, irrespective of the origin of that pattern in the world. In other words, whether the bright–dark boundary is due to the edge of an object, a shadow, a reflection, or a texture marking, if it lines up with the cell's preferred orientation, the cell will still respond. This is important as it means that the cells also tend to respond to the compressed texture elements that occur when texture is projected from a 3D surface into the image. When the surface is slanted, and the texture is compressed in a specific direction, it creates many little streaks in the image, which are perpendicular to the direction of the slant. Cells whose preferred orientation is aligned with the streaks will tend to respond, while those that are tuned to different orientations will respond more weakly or not at all. This is depicted in Figure II.17-5.

As we mentioned earlier, for each location in the retina there are many different cortical cells tuned to all possible orientations. The combined response of all these different cells is referred to as a "population response." When the surface is slanted so that the texture forms streaks at 45° in the image, then the cells that are tuned close to 45° will respond strongly, while the other cells in the population will respond more weakly. As shown in the first panel of Figure II.17-5, this leads to a peak in the population response that is aligned with the direction of the streaks. Importantly, this means that these cells directly measure the direction in which the surface is slanted. In other words, although these cells are often thought of as edge-detectors, they can also play a key role in estimating the 3D shape of textured surfaces. Interestingly, we have found in other research that such cell populations likely play a similar role in estimating shape from shading as well. We next briefly describe how the behavior of the population response provides the brain with information about the surface.

When the slant of the surface changes, the response of each individual cell changes, leading to a concomitant change in the population response. Specifically, as the *direction* of slant changes, the orientation of the streaks in the image also changes, causing different cells to respond most strongly. This causes the peak of the population

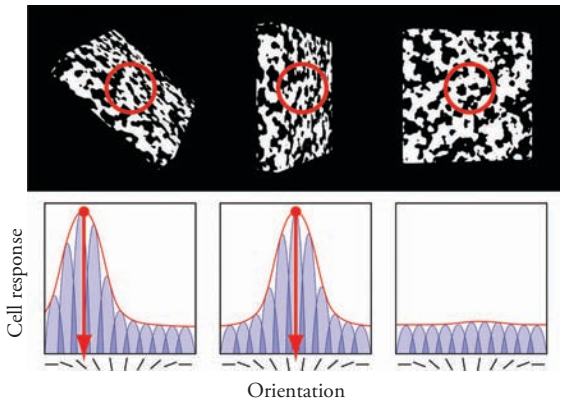


Figure II.17-5. Responses of cell populations to textured surfaces. Blue curves show the orientation tuning functions of individual cells; red curves indicate the population responses to the image patches indicated by the red circles. Left: The surface is slanted so that the texture is compressed at 45° in the image. In the cartoon cell population, the cells tuned close to 45° respond most strongly, while the other cells respond more weakly, leading to a peak in the population response at 45° (red arrow). Middle: The surface is slanted in a different direction, leading to a shift in the peak of the cell population response. Right: The surface is almost frontoparallel, leading to barely any peak in the population response. Images: Roland W Fleming and Daniel Holtmann-Rice.

response to shift, indicating a different surface orientation (Fig. II.17-5, second panel). Additionally, when the *amount* of slant changes, the size of the peak in the population response changes. When the surface is highly slanted, the texture appears extremely streaky in the image, which leads to a strongly defined peak. By contrast, when the surface is close to facing directly toward the viewer, the texture appears hardly compressed at all, so all cells respond with roughly equal strength, leading to a flat population response (Fig. II.17-5, third panel). Thus the population response changes systematically depending on the orientation of the surface, extracting key information about the object's 3D shape.

It is important to note that the cells do not extract individual elements in the texture and work out exactly how they are compressed. The cells do not compute shape using projected rays or the geometry of perspective. They just respond to how streaky or "anisotropic" the texture is at each point in the image. This is why they do not care that the image in Figure II.17-1 is not a physically accurate rendering. The smeared-out texture patterns cause them to respond, and the more smeared out the texture is, the more pronounced the response. This is why the illusion works: it taps into the basic mechanisms that the brain uses to estimate local texture compression.

FUN AND GAMES WITH SHAPE FROM SMEAR

Because we generate shape from smear patterns entirely in 2D, there are a number of tricks we can try to see how they affect the perception of shape. For example, what happens when we increase the amount of smear? As the patterns become more smeared, the texture appears streakier in the image. This means the population response should become more strongly peaked, which should make the surface appear progressively more 3D the more we smear the patterns. The results are shown in Figure II.17-6 (top row). From left to right, we progressively increased the amount by which we smeared the patterns. Most people agree that this makes the patterns appear more and more 3D. This again suggests we have tapped into some crucial mechanisms involved in 3D shape perception. Effectively we have identified which "knobs to turn" to make the illusion stronger or weaker, suggesting we have isolated some key image parameters that the brain uses for estimating shape from texture.

Another trick we can play is to separate out the contributions of blurring and smearing the texture. Recall that in Figure II.17-1 we not only smeared the texture to mimic the directional effects of slant. We also blurred the texture by different amounts at each point to convey the effect of differing amounts of slant on the size of the texture elements. With a physical rendering it is difficult to separate these two effects, but here we can easily enable or disable them independently. In the second row of Figure II.17-6 we show what happens when we use only the blurring (left) or only the smearing (right). Most people agree that the image on the left appears substantially less 3D than the one on the right. In other words, the directional component of the compression seems to be more important than the effect of compression on the size of the texture elements in the image.

Finally, we can try something really weird. In the top row of Figure II.17-6 we progressively changed the amount of smear but kept the directions in which we smeared the texture constant from image to image. This caused the resulting patterns to appear more and more 3D the further we smeared them. But what happens if we progressively change the *directions* in which we smear the noise? Do the patterns still look 3D? This is an interesting question because, according to the theory, the orientations of the texture patterns are meant to have a specific 3D interpretation: they indicate the *direction of the slant* of the surface at each location. However, if we change the directions of the smear at each point in the image (e.g., by rotating them all by 45°), then we create a new pattern, which would indicate surface orientations that are twisted by 45° at every location. The only problem for the brain is that there exists no coherent 3D surface where all the surface orientations are twisted 45°—at least not relative to the shape that we started with. Thus, by changing the smear directions in this way, we create a field of orientation signals that can be measured perfectly well at each location but that make no sense when all the signals are combined across the whole image. No such 3D surface exists, so how is the brain meant to turn the signals into a representation of a 3D object?

In the last row of Figure II.17-6 we show what happens when we rotate all of the directions of smear by a certain angle. On the left, the noise has been smeared along the correct orientations to indicate a specific shape (i.e., no additional rotation was applied to the directions). As we saw earlier, this leads to a compelling impression of a 3D shape. In the middle, we added 45° to all the directions, so that the pattern was "twisted" into a spiral formation. Interestingly, the sense of three-dimensionality started to break down. On the right, we

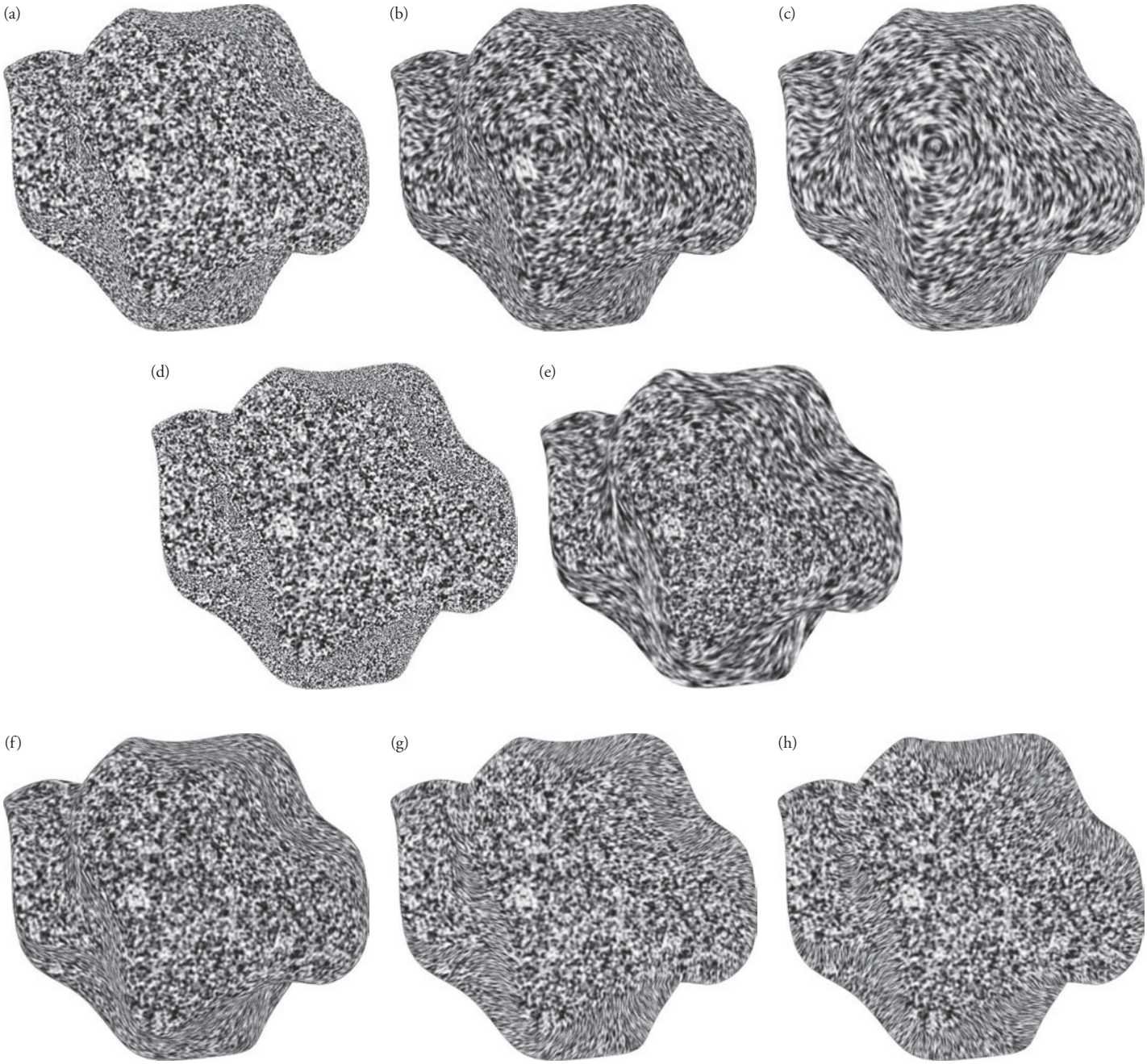


Figure II.17-6. "Turning the knobs" for shape from smear. Top row: Increasing the amount of smear makes the pattern appear progressively more 3D from left to right. Middle row: How important is the directionality of compression due to slant? On the left, what happens when we change the dot sizes in the texture without smearing? On the right, what happens when we smear without changing the dot sizes in the texture? Most people agree the right image appears more 3D than the left image. Bottom row: Changing the direction of smear destroys the perception of a coherent, smooth, 3D surface. Images: Roland W Fleming and Daniel Holtmann-Rice.

rotated all the directions by 90°, so that they are orthogonal to the true directions that would be created by a smooth surface coated in texture. The impression of 3D shape was substantially altered or destroyed. Some people report seeing the pattern as some sort of "fur ball," as if the individual streaks were thin curved strands radiating out of the object. Under this impression, some sense of three-dimensionality is preserved. However, most observers agree that the pattern no longer looks like a single, coherent, smooth 3D surface like the first image in the sequence.

The fact that the perception of 3D shape is radically altered by this simple transformation indicates that the brain uses specific rules for interpreting the patterns of local orientation signals. Importantly, it is not sufficient for a given region of the image to contain oriented patterns. Instead, the orientations have to be systematically organized in to a coherent *orientation field* in order for the brain to interpret them as a surface. The field as a whole must be interpretable as a meaningful surface; otherwise the brain will reject the patterns as having some alternative cause. How exactly the brain combines and groups together signals

from separate locations in the image into a coherent whole is still poorly understood. Hopefully, over the coming years, images like the ones we have presented in this chapter will help us to work out how the brain turns local signals into a representation of a smooth and meaningful 3D surface. Then, and only then, will we fully understand how shape from smear works.

REFERENCES

Fleming, R. W., Holtmann-Rice, D., & Bülthoff, H. H. (2011). Estimation of 3D shape from image orientations. *Proceedings of the National Academy of Sciences USA*, *108*(51), 20438–20443.

Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology*, *148*, 574–591.

Chapter 18
Geometric-Optical Illusions Under Isoluminance?

Kai Hamburger, Thorsten Hansen, and Karl R. Gegenfurtner

THE EFFECTS

This chapter focuses on several illusory effects made up from simple geometric figures and various combinations thereof under conditions of isoluminance (equal luminance; e.g., Liebmann, 1927). We first briefly describe the illusion patterns before we introduce the research and findings under the different luminance conditions.

The so-called geometric-optical illusions demonstrated in this section may be subdivided into three different categories: illusions of size, length, and orientation. Nine replicas of the most famous and well-known of these illusions are presented in Figure II.18-1, namely Delboeuf and Ebbinghaus (size); Judd, Müller-Lyer, Ponzo, and horizontal–vertical (length); and Hering, Poggendorff, and Zöllner (orientation). For a detailed overview of relevant research and literature on these geometric-optical illusions see also Hamburger, Hansen, and Gegenfurtner (2007).

Illusions of Size

Here we normally misinterpret physically equally sized objects as being different in size. This means for the Delboeuf and the Ebbinghaus illusion that the upper inner circles are perceived as being smaller compared to the lower inner circles.

Illusions of Length

These make us erroneously judge various lines of equal physical length as being differently long. In the Judd illusion, this means that even though the dot is located in the center of the horizontal line, the left part of the line is perceived as being shorter than the right part (bisection effect). In the Müller-Lyer illusion the upper horizontal line is judged to be longer compared to the lower horizontal line. In the Ponzo illusion both horizontal lines are equally long but the upper one is perceived as being longer. In the horizontal–vertical illusion the length of the vertical line is overestimated as being longer than the horizontal line.

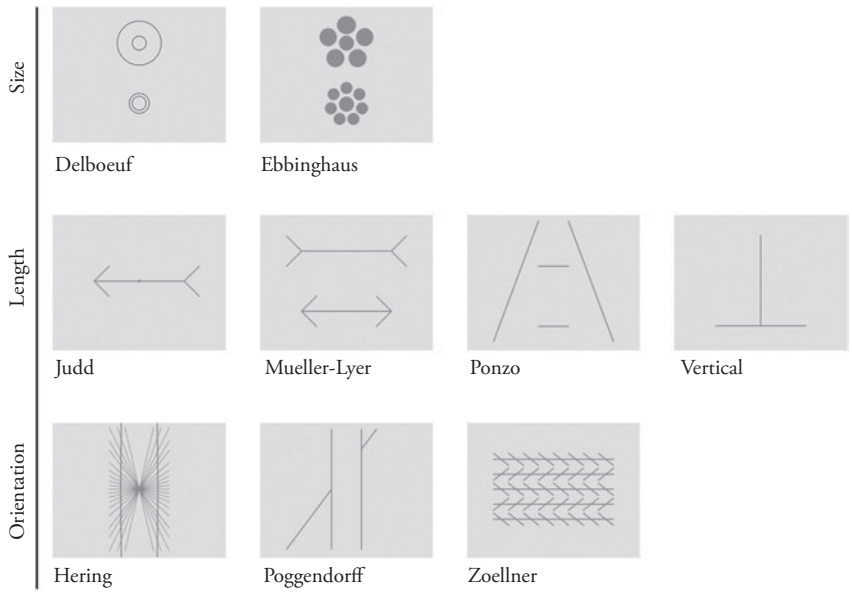


Figure II.18-1. Nine of the most popular and well-known geometric-optical illusions: (top) illusions of size; (middle) illusions of length; and (bottom) illusions of orientation.

Illusions of Orientation

The two straight vertical lines in the Hering illusion seem to be bended, while the disrupted diagonal line in the Poggendorff pattern does not just seem to be a partially occluded diagonal, but rather the upper part seems to be offset. Finally, the physically straight horizontal lines of the Zöllner illusion appear to be tilted and not parallel.

The reader may easily check all of these descriptions of the illusory effects by simply using a ruler and measuring the physical sizes, lengths, and orientations. The knowledge (cognition) about these patterns and the associated effects does not prevent the observer from being tricked by them (perception). In other words, even if we "know," for example, that the two lines in the Müller-Lyer illusion are equally long, we still "see" them as being different. Additionally, some of these effects are reminiscent of real-life situations, such as the fact that the height of a skyscraper appears to be very large, but when the same distance is inspected on the floor (horizontal plane), it apparently appears to be much smaller than in the vertical plane. Context information, namely objects in our immediate surround, also influence judgments of size. This is, for example, the case in the moon illusion: When we look to the horizon in the evening and the moon is very low, then it appears to be rather large, while it appears to be much smaller when it is high in the night sky. However, the retinal image that the moon produces in our eyes is always of the same size. Thus context information at the horizon (e.g., trees, houses) may serve as a comparison that is absent in the sky (there, only stars occur, which are not at all helpful for size comparison).

RELEVANCE

The geometric-optical illusions probably represent the class of visual illusions with the longest history in the vision sciences. They are typically found in almost every textbook, and they serve to demonstrate a variety of visual effects. They are made up from lines and simple geometric shapes and thus are easily constructed and can simply be demonstrated on paper. Another typical characteristic of geometric-optical illusions is that they are mostly presented achromatically (black/white), thus creating a high luminance contrast with the background. These characteristics taken together make them ideal for experiments in vision and cognitive science, since they allow for testing empirical questions on the neurobiology of vision and underlying cognitive processes.

VARIANTS AND PARAMETERS

In order to test for different factors contributing to these illusions, isoluminant variants (no luminance difference between different shapes and background; they just differ in hue) have been tested over several decades (a detailed review is provided by Cavanagh, 1991). Attempts have been made to link the class of geometric-optical illusions to the physiology of early visual processing (e.g., Li & Guo, 1995; Livingstone & Hubel, 1987) providing ambiguous results. Livingstone and Hubel, for instance, claimed that geometric-optical illusions vanish under conditions of isoluminance, while others showed that these effects remain (quite strongly) under such luminance conditions (e.g., Hamburger et al., 2007). Support for the existence of geometric-optical illusions under conditions of equal luminance comes from physiological studies. They have shown that, for example, color and luminance are processed together by the same parvo retinal ganglion cells and neurons in areas V1 and V2 (earliest stages of the visual cortex) that respond to oriented chromatic contrast (for a review see Gegenfurtner, 2003).

A more recent psychophysical study by Hamburger et al. (2007) investigated the nine illusory patterns of Figure II.18-1 under different luminance conditions. The authors used high luminance contrast (black/white), low luminance contrast (10%; shades of grey), and isoluminance along the L–M axis and S–(L + M) axis of DKL color space (Derrington, Krauskopf, & Lennie, 1984), namely red-green and blue-yellow. The isoluminance condition was either a physical isoluminance or a subjective isoluminance level (for which the subjective isoluminance was measured/estimated in advance for each participant). The two (physical) isoluminance conditions are visualized for the Poggendorff illusion in Figure II.18-2 (please note that the actual colors presented here are not isoluminant due to reproduction).

In the experiments participants had to adjust the size, length, or orientation of different components of the patterns to perceptually obtain equal sizes, lengths, or orientations (elimination of illusory effect). If the illusory effects were diminished or even vanished, then participants adjusted the parameters in the opposite direction of the perceived illusory effect (e.g., shortened the upper line in the Müller-Lyer illusion). In all nine patterns the participants compensated for the illusory effect, so that all parameters perceptually appeared as they should, thus introducing physical differences in the stimulus patterns. In general, no significant reductions of illusory strength could be found in these illusions, nor in the low luminance condition or for the two isoluminance conditions. Also, with low-pass filtered edges (blurred edges) the illusions remained stable under all conditions (Hamburger et al., 2007). Thus the results of these controlled experiments challenge the previous findings by Livingstone and Hubel (1987) that geometric-optical illusions are greatly diminished or even absent at isoluminance.

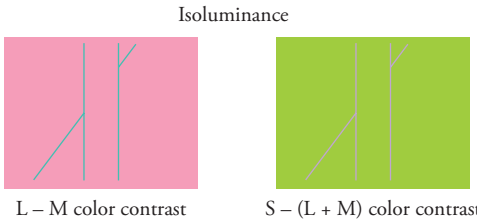


Figure II.18-2. Isoluminance conditions based on DKL color space (Derrington et al., 1984). The lines vary only in hue from the surround while the luminance level is equal. Please note that due to printing the colors may vary in luminance.

CONCLUSIONS

The study by Hamburger et al. (2007) is thus the one (a) with the largest sample of geometric-optical illusions, (b) that first tested under conditions of high luminance contrast, low luminance contrast, and two isoluminance conditions, (c) that first tested under physical and subjective isoluminance, and (d) that first provided a control for chromatic aberrations by additionally using blurred edges (not reported here in detail). In previous research on geometric-optical illusions it has been assumed that they are mediated by the magnocellular system (luminance) and vanish under conditions of isoluminance when only the parvocellular system (color) is activated (Hubel & Livingstone, 1987; Livingstone & Hubel, 1987). This might be the case if processing of different features (color, luminance, shape) in different anatomical pathways were totally independent. But recent research has shown that subcortical pathways combine early in the visual cortex and can therefore not be mapped to independent functional streams (e.g., Lamme & Roelfsema, 2000; Sincich & Horton, 2005). However, Li and Guo (1995) suggested that geometric-optical illusions of size, length, and orientation are mediated by the parvocellular system. This assumption has been put forward by Hamburger and colleagues (2007) by assuming that geometric-optical illusions are not only mediated by the parvocellular system but also by the koniocellular system. A possible explanation for why some studies (e.g., Liebmann, 1927) found a breakdown of geometric-optical illusions at isoluminance may be a result of lines that were too thin to be resolved by the chromatic system.

This chapter demonstrates that geometric-optical illusions do not disappear under conditions of low luminance contrast or different conditions of isoluminance. On the contrary, there is now much evidence available that they remain stable with variations of luminance contrast. This is consistent with the assumption of joint processing of oriented color and luminance contrast in early visual areas. Finally, our conscious percept of geometric-optical illusions is similarly affected for luminance conditions and isoluminance. This suggests that joint processing of chromatic- and luminance-defined contours extends well beyond early visual areas.

REFERENCES

Cavanagh, P. (1991). Vision at equiluminance. In J. J. Kulikowski, I. J. Murray, & V. Walsh (Eds.), *Vision and visual dysfunction: Limits of vision* (Vol. V, pp. 234–250). Boca Raton, FL: CRC Press.

Derrington, A. M., Krauskopf, J., & Lennie, P. (1984). Chromatic mechanisms in the lateral geniculate nucleus of macaque. *The Journal of Physiology, 357*, 241–265.

Gegenfurtner, K. R. (2003). Cortical mechanisms of color vision. *Nature Reviews Neuroscience, 4,* 563–572.

Hamburger, K., Hansen, T., & Gegenfurtner, K. R. (2007). Geometric-optical illusions at isoluminance. *Vision Research, 47*, 3276–3285.

Hubel, D. H., & Livingstone, M. S. (1987). Segregation of form, color, and stereopsis in primate area 18. *Journal of Neuroscience, 7*, 3378–3415.

Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neuroscience, 23*, 571–579.

Li, C.-Y., & Guo, K. (1995). Measurement of geometric illusions, illusory contours and stereo-depth at luminance and colour contrast. *Vision Research, 35*, 1713–1720.

Liebmann, S. (1927). "Über das Verhalten farbiger Formen bei Helligkeitsgleichheit von Figur und Grund." *Physiologische Forschung, 9*, 300–353.

Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience, 7*, 3416–3468.

Sincich, L. C., & Horton, J. C. (2005). The circuitry of V1 and V2: Integration of color, form and motion. *Annual Reviews in Neuroscience, 28*, 303–328.

Chapter 19
The Picture Surface Illusion

3D Biases 2D

Sherief Hammad and John M. Kennedy

INTRODUCTION

The leftmost drawing in Figure II.19-1 is based on one of a cube balanced on a point (after Kennedy & Juricevic, 2006). Its top quadrilateral shows a tilted square. The quadrilateral on the picture surface appears to be a rectangle, but its angles are actually obtuse and acute. In fact they differ by more than 20˚. The eye is tricked. The 2D angles are seen biased toward the depicted three-dimensional (3D) angle, 90˚ (Hammad, Kennedy, Juricevic, & Rajani, 2008b). In some versions of this illusion, the combined error in reading the two angles is 40°. It is also a "double-cube" in parallel perspective. The top's diagonal line would create angles in the 3D double-cube of 27˚ and 63˚. The two-dimensional (2D) angles on the picture surface are seen biased toward these values. The trio on the right shows cubes in linear perspective, tilted. Shown statically, at tilts of 25˚ to 65˚ the top quadrilateral's left angle is misperceived by about 15˚ and the obtuse angle by 10˚ to 30˚ degrees. Forward and back tilting of the cubes increases the error.

The double-cube picture is in parallel perspective. The cubes are tilted with respect to the picture surface and are shown foreshortened so the top quadrilaterals are "eccentric" (i.e., the horizontal dimensions are larger than the vertical). The three cubes on the right in Figure II.19-1 are in linear perspective, which uses convergence as well as eccentricity. In the top quadrilaterals pairs of lines converge upward, symmetrically to the left and right, and pairs of lines for the flanks of each of the cubes converge downward. This three-point perspective reveals 3D space. Also, angles embraced by the top L-junction and the Y-junction below it change as the cube tilts forward.

In linear perspective, if the cube is tilted at 45˚ to the picture surface, foreshortening produces a height/width aspect ratio (eccentricity) of the top quadrilateral of .65. If the cube top is tilted closer to parallel to the picture surface, with the top face of the cube tilted to 65˚ from the horizontal, the aspect ratio is .86. Tilted 15˚ it is .24. Also, each cube projects a 2D Y-junction. On its own, this looks like a flat form, but in Figure II.19-1 it suggests a 3D corner, with distance from the Y indicating depth. Other vertices of the cubes are shown by arrowheads and L-junctions. If enough detail is present (Gibson, 1950), the figures function as pictures of cubes: the eccentricities become foreshortenings and the junctions stand for vertices of 3D blocks and the lines for edges of surfaces (Kennedy & Hammad, 2010; Maniatis, 2009a, 2009b; Pizlo, 2008).

In perception, information for a picture surface flattens pictured-world 3D shapes (Hecht et al., 2003; Koenderink & van Doorn, 2003; Sedgwick, 2003). Pictured 3D features are seen "flattened" as if biased by the 2D surface. Picture surface illusions are the converse—perceived 2D is biased by the picture's 3D information. This is not the same effect as a distant object's apparent size looking small, biased because it subtends a tiny angle (Arnheim, 1954; Kennedy, Juricevic, Hammad, & Rajani, 2007). Rather, picture surface illusions concern features on a surface, sizable visual angles, and 2D angles that look larger than is true. More relevant is the Gregory (1963) proposal that the Müeller-Lyer illusion is caused by 2D arrowheads and Y-junctions looking like corners in 3D. But the arrowheads and Ys can be replaced by squares and the Müeller-Lyer illusion remains. Basically, Gregory had the right idea but the wrong application.

The picture surface illusion may be caused by vision combining (a) information for a flat 2D surface provided by lines, junctions, acute angles, and obtuse angles and (b) 3D pictorial depth and tilt information (found in *perspective* features, notably eccentricity and convergence, and in *configurations* of lines, junctions, and angles that depict 3D objects). If so the illusion will be (a) maximal when *tilt-in-depth (foreshortening) information* is most effective (tilt information can be obvious in pictures of many kinds of objects; see Fig. II.19-2), so (b) the illusion should also occur in pictures of tall and short blocks (Fig. II.19-2, top row) not just cubes. Indeed, (c) it could occur in pictures of tilted irregular pyramids (Fig. II.19-2, middle left). It could increase if compelling (d) texture (especially in perspective) and (e) motion information for pictorial depth is added, perhaps even if the information is added to a single quadrilateral, and (f) decrease if texture emphasizing the 2D picture surface is added (Hammad, 2014, added "paper"). Many shapes can be tilted with respect to the picture surface, so the illusion could occur (g) for ellipses depicting tops of cylinders (Fig. II.19-2, lower right). If eccentricity and the shapes of junctions are adequate information for

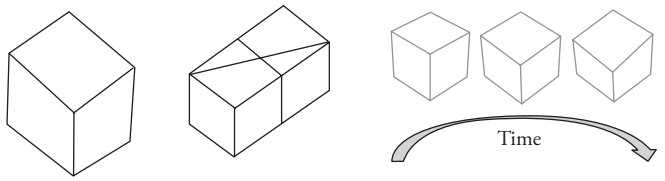


Figure II.19-1. Far left is based on a rough sketch of a cube (Kennedy & Juricevic, 2006). Beside it is a "double-cube" in parallel perspective. The trio on the right shows cubes in linear perspective, tilted forward and backward.

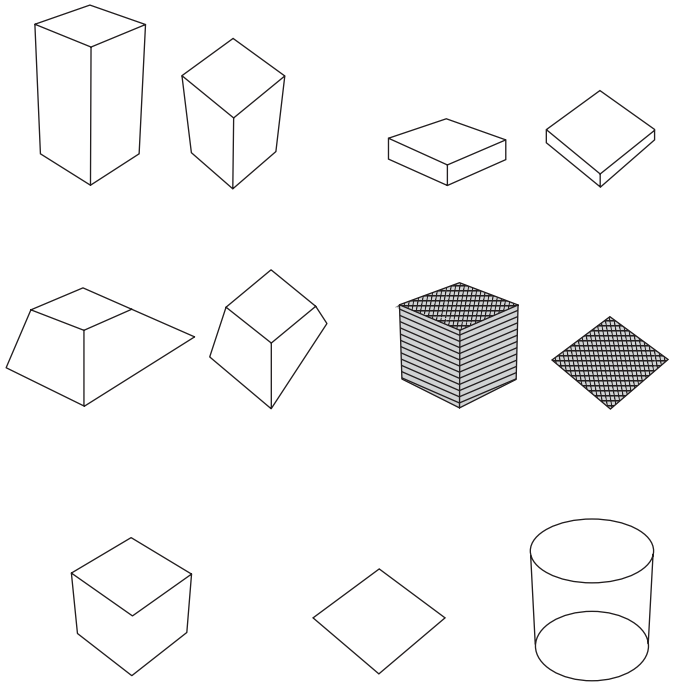


Figure II.19-2. The illusion could be present in pictures of short and tall blocks, irregular pyramids, cylinders, and textured forms but less so in a picture with no Y or L junctions or a textured surface.

tilt, and convergence is not essential, the illusion could be as large in (h) parallel perspective as in linear perspective pictures (Maniatis, 2009a, 2009b). In the absence of arrowheads and Y-junctions, the illusion may be quite slight in an untextured quadrilateral (Fig. II.19-2, lower middle).

PICTURE SURFACE ILLUSION: PARAMETERS

The tilt of linear-perspective cubes was varied as shown in the three cubes in Figure II.19-1, right (Hammad & Kennedy, 2008a). Each cube was static, depicted in linear perspective with white lines on an otherwise black monitor screen, and viewed binocularly. (The observers in studies reported here were University of Toronto undergraduates with normal or corrected-to-normal vision, and the experiments involved 8 to 16 observers.)

At one cube tilt (5˚ from horizontal, 85˚ away from a vertical picture surface), the top quadrilateral is thin (eccentricity 1 to 14) and the cube flanks are shown by lines that are nearly parallel. The depicted object may look flat—or "squashed" (Maniatis, 2009a). At 85˚ from horizontal, 5˚ away from the picture surface, the quadrilateral is almost a square (proportions 1 to 1.01), the lines for the flanks are miniscule, and the cube top is almost parallel to the picture surface, not slanted into depth. Also, a quadrilateral on its own may seem largely 2D, slanted very little if at all. Ergo a slanted surface may be most obvious at intermediate tilts of the cube.

Observers, seated 75 cm from the screen, the distance of the correct center of projection (Todorovic, 2009), with cube pictures subtending about 7˚, judged the size of the top obtuse angle and the leftmost acute angle in the quadrilaterals depicting the top of the cubes and in quadrilaterals shown on their own, with no lines for the flanks of the cube. Tilts from 5˚ to 85˚ from horizontal were tested in 10˚ steps, resulting in acute angles on the screen, presented to observers, of 9˚, 26˚, 41˚, 54˚, 65˚, 75˚, 80˚, 87˚, and 89˚. The obtuse angles were 171˚, 157˚, 143˚, 130˚, 120˚, 109˚, 103˚, 96˚, and 91˚. (The upper and lower obtuse angles are not equal in a linear-perspective projection, so the sum of an acute and obtuse angle is not always 180˚.)

A 2D angle of 41˚ (with cube tilt 25˚) was reported to be 57˚, an error of 16˚, or almost 40% (Fig. II.19-3). The 2D 54˚ (at tilt 35˚) was seen as 69˚, an error of 15˚. Also, 65˚ (at tilt 45˚) was reported as 79˚, a 14˚error. In contrast, the error in viewing the quadrilateral on its own, with no lines for the flanks of the cube, was slight. The mean absolute error in judging the acute angle in the quadrilateral on its own was 2˚. Though its 130˚ obtuse angle was underestimated as much as 10˚, this is less than 10% of the obtuse angle. The quadrilateral offered only its aspect ratio and angles at L-vertices as tilt information, representing slant information that is meager and unlikely to be effective.

The error in judging the cube-picture obtuse angle was greatest for angle 157˚ (tilt 75˚), an error of 31˚, so the mean reported angle was 136˚. Since the error in judging the acute angle accompanying the 157˚ obtuse angle was 9˚, the difference between the 2D angles was reduced in perception by 40˚. The largest errors in judging the obtuse angle were from 120˚ (error 18˚) to 157˚ (error 31˚)—that is, at the intermediate tilts.

At extreme tilts the errors may be subject to ceiling and floor effects. Small 2D angles might have relatively small errors, which are hard to measure. If the bias is toward 90˚, errors for 2D angles near 90˚ are also hard to measure.

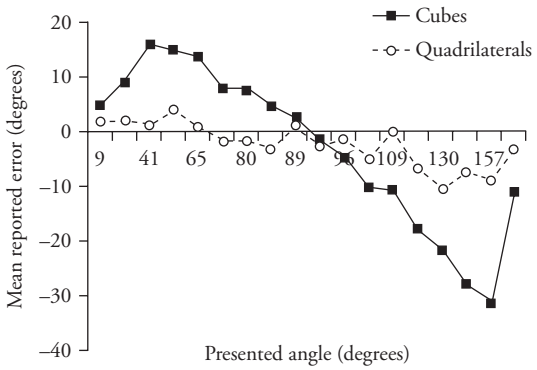


Figure II.19-3. Mean response errors in degrees in judging 2D acute and obtuse angles in pictures of tilted cubes and slanted squares.

However, reduction of large obtuse angles escapes floor and ceiling problems. If so, the fact that the error for the largest obtuse angle is small may be due to the configuration appearing like a flat object, or "squashed" (Maniatis, 2009a), not slanted in depth.

If the prominence of errors at intermediate tilts is partly due to the ceiling and floor effects at other tilts, this is hardly compelling evidence for the tilt-in-depth account of the illusion. So let us pursue the matter further. An important factor in creating the illusion may be the tilt information present in the top quadrilateral, rather than the converging lines for the cube's flanks. If so, short 2D lines that are just long enough to suggest the quadrilateral is thick and not lying flat on the picture surface may suffice. In a second study, all the cubes were replaced by tall and short parallelepipeds and flanks tall as double-cubes (tall blocks) or short as quarter-cubes (resembling tiles), as in Figure II.19-2. Observers judged the top left acute angle. Errors in both cases were not significantly different from those with cubes—about 15˚ degrees for 25˚ to 45˚ acute angles—suggesting that any lines for cube flanks, whether short or long, may allow information for 3D tilt in the 2D quadrilateral to become effective in perception (Hammad, 2014).

If tilt information in the 2D quadrilateral is key, the illusion might not require familiar block objects and all angles be right angles. Lines suggesting the flanks of pyramids might be enough. A third study tested this with pictures of irregular pyramids (Fig. II.19-2, middle row, two left figures). The flanks of the pyramids were shown with lines of different lengths, converging upward in the picture. The lines were shorter than the Y's central long line. All of the upward-converging lines were drawn at different angles to the vertical. They were lower than the quadrilateral's top two lines. In effect, the quadrilateral represented the top face of an irregular flat-topped pyramid. The pyramid at each tilt was unique for each observer, and no pyramid was presented twice across observers.

Each pyramid can readily be taken to be an irregular 3D block object, with the Y-junction depicting a 3D convex vertex that is "bulky" (its enclosed angles are larger than right angles). The top quadrilateral offers aspect ratios, convergence, and L-junctions of varied sizes—or perspective information for a square at a specific slant. The result? Errors in judging the acute angle in the pyramid's quadrilateral were almost indistinguishable from the errors in the cube experiments. If anything, the maximum error was slightly elevated—18˚ in judging acute angles of 41˚, 2˚ more than for cube pictures.

Note that in a variation of this study, the top quadrilateral was reduced to an isosceles triangle, meaning the stimulus was just the two lower lines of the top quadrilateral and a line joining the left and right vertices. As a percentage of the angle being estimated, the illusion was only slightly reduced (e.g., from 26% at 25˚ tilt to 21%). This is remarkable, since a triangle does not have the asymmetry and eccentricity that allows a foreshortened square to be suggested by a quadrilateral. Nonetheless, the perceiver may take the upper line of the triangle as being to the rear, the rest of the triangle as showing a surface projecting forward, and the isosceles shape as having the aspect ratio of half-a-square, tilted.

Texture is information for a continuous surface and effective as slant-in-depth information. Addition of a texture gradient to the quadrilateral could make its aspect ratio compelling information for slant. A gradient in the texture could boost any modest illusion generated by the quadrilateral on its own and enhance the strong illusion in the cube pictures. In a fourth study, each face of the 3D cube, and the corresponding 3D cube top face on its own, was given a texture of 15 • 15 small squares (Fig. II.19-2, middle row right). As a result, they projected texture gradients on the display screen.

The illusion increased. For tilts of 5˚ to 65˚ the textured cube generated higher errors than the untextured one. For 2D angles between 26˚ and 54˚, the textured cube generated an error a little over 20˚, rather than the maximum of 16˚ with the untextured cube. The error for the textured top face on its own was 10˚ to15˚ for acute angles of 26˚ to 55˚, rather than the mean error of 2˚ as before. (Similar effects occur for parallel-perspective projections of textured cubes and quadrilaterals.)

Like texture, motion is compelling spatial information and includes kinetic depth effects (Wallach & O'Connell, 1953). If centered on its front vertex, a cube tilts forward and back as indicated by the time arrow under the trio of cubes on the right of Figure II.19-1. The top quadrilateral's 2D angles are projected as expanding and contracting. Its lines converge and diverge. The result is information for a cube in motion, and the tilting-in-depth information could result in especially large illusory effects.

In a fifth study, the cubes were shown at 15 frames/s rotating 3˚ around the horizontal axis. The cube top became taller, with larger acute angles, then shorter, then taller and shorter again before returning to its original size and being removed from the screen. The result was that acute angles in the 2D display of 26˚ to 54˚ were reported 22˚ to 23˚ larger (Hammad, 2014). Further, in a similar study with tilting quadrilaterals, in which observers judged the smallest angle presented, error in judging 2D angles of 31˚ and 59˚ were 13˚ and 11˚, not far from the maximum error for static cube pictures.

If adding 3D information increases the illusion, adding information for the 2D surface might lessen it, a conjecture tested in a sixth study. In one condition, the picture surface was made white (with the 2D lines to be judged black) so the monitor surface might be more evident than would be the case for a black screen (with the 2D lines white). In a second, black vertical lines were added to the white screen, simulating lined paper. Adding 2D texture such as parallel 2D verticals could facilitate judging 2D angles (Hammad, 2014, used a detailed scan of lined paper).

Indeed, the illusion diminished. The maximum error was lessened by the white screen to 10˚ at acute angle 41˚ and lessened still further in the white screen plus vertical lines condition to 6˚ in the lined-paper simulation, compared to 16˚ in the black screen condition. The reductions were evident at all of the intermediate tilts tested.

If the illusion relates to tilt-in-depth information, it should apply to an ellipse on the top of a cylinder (Fig. II.19-2 bottom row, right). To test this prediction, in a seventh study ellipses depicting tops of tilted cylinders were matched by observers with a set of ellipses of various eccentricities. Indeed, the resulting matches were less eccentric than veridical (Hammad, Kennedy, Juricevic, & Rajani, 2008a), looking fatter than was true, especially in

the midrange of tilts of the cylinders. Similarly, the width and height diagonals in the top quadrilateral in Figure II.19-1 (left) look equal, but actually the height diagonal is shorter (Shepard, 1981).

Parallel perspective does not use convergence—it only foreshortens. However, for compact objects like cubes subtending small angles such as 7˚ at the observer's vantage point, linear perspective's convergence is often insignificant. Hence parallel-perspective versions of the figures should produce the illusion (Maniatis, 2009b). Indeed, in an eighth study, linear and parallel projections of cubes produced virtually the same errors, with a maximum error of 13˚ to 16˚ at tilts of 15˚ to 35˚.

DISCUSSION

A picture's depth information biases perception of its 2D pictorial features, including straight line and curved shapes, and distances as well as angles. Two-dimensional aspect ratios may be an effective source of 3D slant information, since aspect ratios can be similar in parallel and linear perspective, and they produce similar biases. Also, in depicting the flanks of 3D blocks, short lines whose convergence is so hard to discern they may as well be parallel-produced errors similar to long lines that converge fairly obviously.

The quadrilateral on its own has the same eccentricity as a picture of a cube top but triggers small errors. Nevertheless, textured quadrilaterals had errors close to those of the untextured cube. The texture took the form of a gradient projected from 15 • 15 tiny squares. The gradient likely made eccentricity compelling as tilt information. To-and-fro tilting also provided perspective changes in the quadrilateral, in sequence, and made for compelling tilt information. (Since texture can be sparse or dense, and motion great or small, any difference between the texture- and motion-driven illusions is a comparison of apples to oranges.) Evidently to make a quadrilateral's eccentricity and L-junctions compelling as information for tilt-in-depth, something needs to be added. Otherwise it may be taken as flush with the picture surface.

In the cube picture, eccentricity, Ys, arrowheads, and L's are information for tilt. With tilt, Y's become more like T's, and in the limit arrowheads become L's and L's in the quadrilateral become single lines. Y and arrowhead shapes may have fostered biases in a picture with an isosceles triangle as the top of a pyramid; the triangle may have been taken as depicting half-a-square.

Different features may combine their influence to increase bias. The maximum error for the acute angle in the untextured, static cube case was 16˚ and for the textured quadrilateral was 15˚. Combined, should the sum be 31˚ for the textured cube? Alas, its maximum was barely 20˚. The cube and the texture share common elements, and each may bring only a little unique information to the combination. The unique features could be the Y in the cube and the X junctions in the texture. Perhaps the formula is as follows: common elements (eccentricity generating 10˚ error) + the cube's unique features such as Y's (6˚ error) + the texture's unique features (X's, 4˚ error). This would be perfectly additive: 20˚.

An additive effect is the hallmark of an illusion. Perception is unable to treat different sources of information in an illusory figure as distinct. In contrast, in realistic use of information, one source of information simply confirms another. Each allows cross-checks. Stereo, motion, and texture can all specify a lion is dangerously close, say at 100 m distance. Perception should not add the three channels of information and see the target as safely 300 m away. Better yet, perception should treat each channel's conclusion separately and seek reasons for any discrepancy.

CONCLUSIONS

Drawing is likely to be affected by the picture surface illusion, especially in copying a picture. As we surround a V shape in our drawing with features suggesting 3D, our V shape will change its apparent shape.

Consider viewing a photo of a landscape with a billboard. If the billboard depicts a cube, and the billboard was photographed at a glancing angle, the cube can appear like a ship with a sharp prow (Coren, personal communication, April 6, 2011; Parks, 2001; Pirenne, 1970, p. 97). If our account is correct, the angle on the picture surface should be seen biased toward the 3D angle of the sharp appearance.

Anamorphics are pictures designed to be viewed from glancing angles. Perception of their 2D features should be biased toward their apparent, depicted 3D features. Conversely, the periphery of a wide-angle perspective picture has anamorphic effects (Kubovy, 1986, pp. 107–108). Viewed from distances other than the center of projection, they appear to show highly skewed 3D objects. The skewed 3D angles could be reflected in observations of 2D angles on the picture surface.

One factor remains obscure: how to predict the size of the illusion. Why 10˚ or 20˚, exactly? This is still puzzling. We conclude that information for surfaces slanted in depth, for example eccentricity indicating foreshortening, biases perception of 2D features on picture surfaces.

ACKNOWLEDGMENTS

We thank Lydia Maniatis, Peter Coppin, Marta Wnuczko, Marcelo Santos, Karan Singh, Matthias Niemeier, and George Cree for comments. We also offer our appreciation to Sandacre Technology for programming.

REFERENCES

Arnheim, R. (1954). *Art and visual perception*. University of California Press: Berkeley.

Gibson, J. J. (1950). *The perception of the visual world*. Boston: Houghton-Mifflin.

Gregory, R. L. (l963). Distortion of visual space as inappropriate constancy scaling. *Nature, 199*, 678–691.

Hammad, S. (2014). *Components of optical information for depth induce a picture-surface angle illusion* (Unpublished doctoral dissertation). University of Toronto, Toronto, Ontario.

Hammad, S., Kennedy, J. M., Juricevic, I., & Rajani, S. (2008a). Ellipses on the surface of a picture. *Perception, 37,* 504–510.

Hammad, S., Kennedy, J. M., Juricevic, I., & Rajani, S. (2008b). Angle illusion on a picture's surface. *Spatial Vision, 21*, 451–462.

Hecht, H., Schwartz, R., & Atherton, M. (Eds.). (2003). *Looking into pictures: An interdisciplinary approach to pictorial space*. Cambridge, MA: MIT Press.

Kennedy, J. M., & Hammad, S. (2010). Maniatis's arguments about impossibilities and rigging in Hammad et al. (2008): The importance of Pizlo (2008). *Seeing and Perceiving*, *23*, 75–80.

Kennedy, J. M., & Juricevic, I. (2006). Foreshortening, convergence and drawings from a blind adult. *Perception*, *35*, 847–851.

Kennedy, J. M., Juricevic, I., Hammad, S., & Rajani, S. (2007). Perspective: Arnheim and in-between solutions. *Psychology of Aesthetics, Creativity and the Arts, 1*, 40–46.

Koenderink, J. J., & van Doorn, A. J. (2003). Pictorial space. In H. Hecht, R. Schwartz, & M. Atherton (Eds.), *Looking into pictures: An interdisciplinary approach to pictorial space* (pp. 239–299). Cambridge, MA: MIT Press.

Kubovy, M. (1986). *The psychology of perspective and Renaissance art*. Cambridge, UK: Cambridge University Press.

Maniatis, L. M. (2009a). Comment on "Angle illusion on a picture's surface" by Hammad *et al*. (2008). *Spatial Vision*, *22*, 555–558.

Maniatis, L. M. (2009b). Letter to the editor. *Perception*, *38*, 1416–1419.

Parks, T. E. (Ed.). (2001). *Looking at looking*. London: SAGE.

Pirenne, M. H. (1970). *Optics, painting & photography*. Cambridge, UK: Cambridge University Press.

Pizlo, Z. (2008). *3D shape: Its unique place in visual perception*. Cambridge, MA: MIT Press.

Sedgwick, H. A. (2003). Relating direct and indirect perception of spatial layout. In H. Hecht, R. Schwartz, & M. Atherton (Eds.), *Looking into pictures: An interdisciplinary approach to pictorial space* (pp. 61–76). Cambridge, MA: MIT Press.

Shepard, R. N. (1981). Psychological complementarity. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 279–342). Hillsdale, NJ: Lawrence Erlbaum.

Todorović, D. (2009). The effect of the observer vantage point on perceived distortions in linear perspective images. *Attention, Perception & Psychophysics*, *71*, 183–193.

Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, *45*, 205–217.

Chapter 20
Cast Shadow Illusions

Daniel Kersten and Pascal Mamassian

INTRODUCTION

Shadows are everywhere, but they usually go unnoticed. We do, however, use their consequences all the time. Figure II.20-1A shows two pictures, each with a green square in front of a checkerboard background. The two pictures differ only in the cast shadows of the square. Because of these differences, the right square appears to be further from the background than the left square. By adding motion to the cast shadow, the effect on the perception of depth can be made quite striking. Moving the shadow, leaving the square fixed in the image, results in the square appearing to move in depth (see the Companion Website http://youtube/ig2CV1TLn5A). It is only with scrutiny that one is convinced that the square is not moving at all within the picture frame. In particular, the image of the square is not changing size, as it should if it was indeed moving in depth. Shadows seem to create a paradox for the visual system. While on the one hand they can be quite useful as cues for depth, their unobtrusiveness suggests that there may be good reasons for a visual system to get rid of them.[1]

SHADOWS: USE THEM OR LOSE THEM?

As visual animals, we typically need information about objects for the job at hand, whether that be describing, handling, navigating around, or tracking them. It would make sense that shadows go unnoticed because our attention is primarily directed toward the tangible objects we need rather than toward their shadows. Depending on the task, it might even make sense for the visual system to filter out shadows completely—for example, when they fall on the very object of interest, making it difficult to detect the features specific to the object (see Fig. II.20-2a; Cavanagh 1991; Lovell, Gilchrist, Tolhurst, & Troscianko, 2009). The problem of detecting object-specific features is part of a more general problem that the brain must solve to reliably recognize objects. The images of an object (called its "appearances") vary greatly from one instance to the next because of changes in both viewpoint and illumination. Illumination gives rise to shading patterns that are affected by the object's orientation relative to the illumination sources and the viewpoint. And shadows, in particular, can have a strong effect on appearances.

Given the confounding nature of shadows, it may not be surprising that we seldom notice shadow inconsistencies in paintings or when artificially created in photographs (Cavanagh, 2005; Mamassian, 2004; Ostrovsky, Cavanagh, & Sinha, 2005). Even with careful inspection we are not very good at deciding whether a cast shadow is faithful to the laws of optics or is faked (e.g., see the analysis of the Lee Harvey Oswald backyard pictures; Farid, 2009; Fig. II.20-2b). Further, pictorial artists often eliminate shadows that might interfere with interpretation (Gombrich, 1995; Jacobson & Werner, 2004). Human visual recognition can suffer in the presence of shadows (Braje, Kersten, Tarr, & Troje, 1998; Braje, Legge, & Kersten, 2000; Tarr, Kersten, & Bülthoff, 1998), and considerable computer vision research has gone into detecting shadows in order to remove them (Finlayson, Hordley, & Drew, 2006; Jacobson & Werner, 2004). Does their confounding effect on recognition (Porter, Tales, & Leonards, 2010), together with the need to attend to objects, imply that shadows are more trouble than they are worth and that the visual system filters them out? Our illustration of depth from cast shadows suggests that is obviously not the case. In fact, artists, photographers, animators, and lighting designers have long known that shadows have positive as well as negative consequences for perception—their judicious placement adds a sense of depth and realism and evokes mood, whereas injudicious placement of even optically correct shadows can be misleading or distracting. So while shadows do not have a starring role to play on the stage of conscious perception, the illusory motion demonstration described earlier does suggest that they have a strong supporting role in conveying information about depth. What kinds of parts do shadows play in depth perception, and just how strong are their roles? What information does the human visual system extract in order to use shadows?

THE PERCEPTUAL USE OF SHADOWS

Because of their role in providing depth information, shadows can be useful for diverse tasks such as deciding whether two surfaces are touching (Thompson, Shirley, Smits, Kersten, & Madison, 1998), estimating the future trajectory of a flying object (Taya & Miura, 2010), determining the structural form of patterns (Khuu, Khambiye,

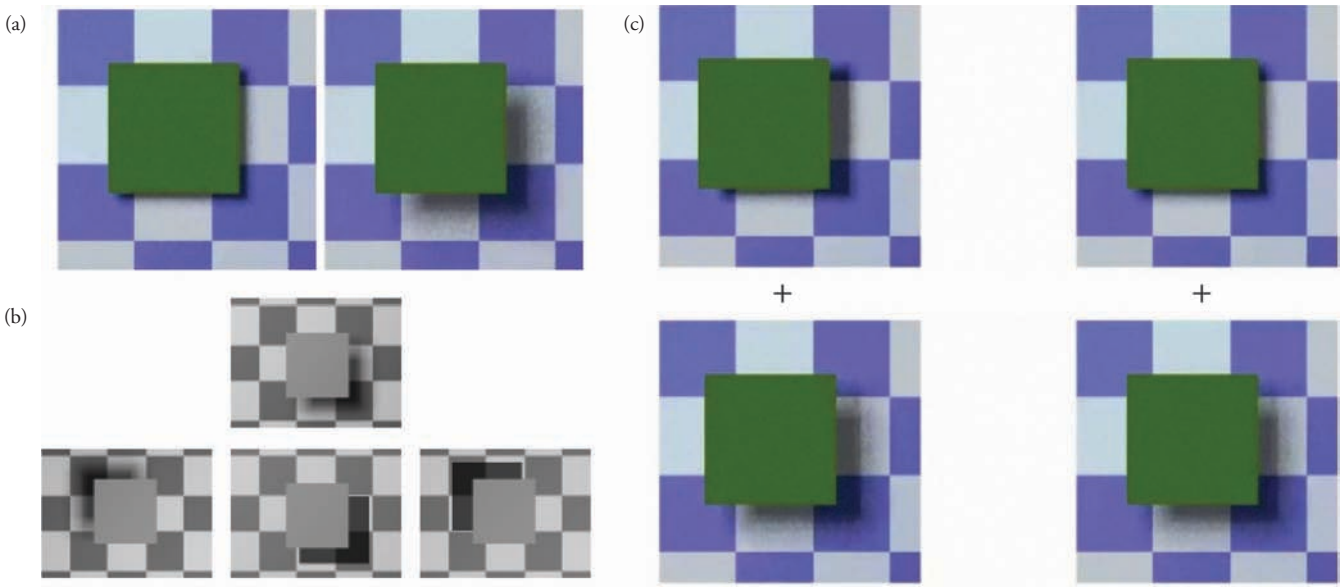


Figure II.20-1. (a) A cast shadow provides useful information for the depth of an object relative to the background surface. One can produce the appearance of motion in depth of the unmoving square by moving the shadow over a sequence of frames (see the Companion Website http://youtube/ig2CV1TLn5A). (b) Cast shadows are most effective at conveying depth when they have typical properties (top panel). If the shadows are above the object, or are sharp, they tend to be less effective (bottom three panels; Kersten et al., 1996). (c) Information about depth from shadows can be placed in conflict with information from stereo disparity. The stereo disparity provides information indicating that the top square is further from the background, whereas the shadow indicates it is the bottom square that is further away from the background. There are differences between people, and readers may be able to judge for themselves which cue wins. To experience the stereo effect, try to cross the two eyes so that the left and right images are seen by the right and left eyes, respectively. An animated version of this demonstration, with moving shadows, can be viewed at the Companion Website: http://youtube/d4KMLLSs1C0.

& Phu, 2012), deciding whether one surface occludes another (Tomonaga & Imura, 2010), and detecting differences in shape (Rensink & Cavanagh, 2004). Infants as young as five to seven months use cast shadows for depth (Imura et al., 2006; Yonas & Granrud, 2006). Cast shadows improve perception of relative depth in chimpanzees (Imura & Tomonaga, 2009). Shadow features in an image, unlike specular highlights, can be directly tied to surface points



Figure II.20-2. (a) Shadows cast from other objects create problems for visual recognition. Upper and lower panels show the original gray level image and the output of a standard edge detector, respectively. While cast shadow edges can, in theory, provide some information regarding shape, they tend to be more trouble than they are worth. (b) It is not easy to tell whether cast shadows are physically correct, as illustrated by a photograph of Lee Harvey Oswald, the alleged assassin of John F. Kennedy (*LIFE Magazine*, February 21, 1964). Conspiracy theorists have argued that Oswald's claims that the photographs were doctored were supported by the supposed "fact" that shadows cast by his nose and body, were not consistent with a single light source (Farid, 2009).

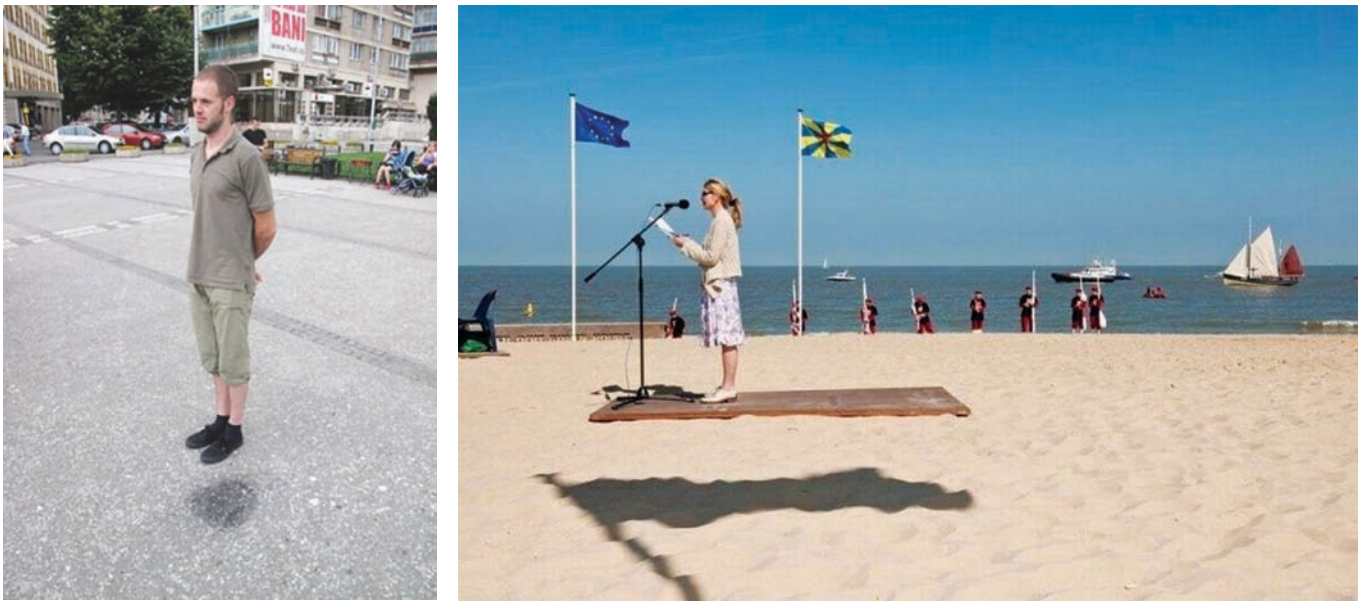


Figure II.20-3. Left panel: The absence of a strong true shadow of the man, together with a "good-enough" substitute—the dark spot on the road—produces apparent levitation. Right panel: The shadow of a flag appears to belong to the platform, creating the illusion that the speaker is elevated with no visible means of support.

for stereo disparity, which in turn provides information for three-dimensional shape (Puerta, 1989). (For a general review of the role of cast shadows in perception, see Dee & Santos, 2011. For a historical and cultural treatment of shadows, see Casati, 2004a.)

The role of cast shadows as a means to convey depth in drawings was recognized long ago by Leonardo da Vinci (Fiorani, 2008). One aspect of his advice—the ability to add "drop shadows"—is used every day in graphics presentation and layout software (e.g., as a text format option in Microsoft's PowerPoint and Apple's Keynote presentation software). The strength of real-life drop shadows is illustrated below by putting them in conflict with our daily experience that objects should not defy gravity.

How Strong Are Cast Shadow Cues for Depth?

A critical visual function is to know where an object is on a ground plane. Cast shadows provide pervasive and strong cues for contact between one surface and another, acting as a kind of "visual glue" (Thompson et al., 1998). The information lies in the proximity of the image of an object and its cast shadow and the sharpness of the penumbra, together with how local shadow and object contours meet (Madison, Thompson, Kersten, Shirley, & Smits, 2001). When there is no visual glue, the depth of an object becomes ambiguous. As an object rises above a plane, the shadow gets further from the casting object, and the size of the penumbra increases.[2]

Apparent Levitation

Cast shadows are sufficiently powerful as a depth cue; that lack of contact information can inveigle perception to believing that objects or people can defy gravity, a trick used by street magicians. Here the magician raises one foot, creating a displaced cast shadow while using the same foot to hide the fact that the other foot is holding his body up tiptoed (cf. see the Companion Website https://en.wikipedia.org/wiki/Balducci_levitation).

Another trick is to take advantage of a gray day when the contact cue is weak, increasing depth ambiguity. Viewing from just the right angle allows a distant wet spot to assume the role of cast shadow, resulting in apparent levitation (Fig. II.20-3, left panel). Even on a bright day, an overhead sun can make an object's true shadow hard to see, reducing its role as visual glue, allowing a visually more salient shadow from another object to become associated with the object of interest, and making it appear to float mysteriously above the ground (Fig. II.20-3, right panel).

By creating variations of the green-square animation described earlier, one can show that the strength of the apparent change in depth is stronger if (a) the light source appears to come from above, (b) the cast shadow has a visible penumbra whose fuzziness grows with relative depth, and (c) the shadow is dark rather than light (which is physically impossible). These factors all suggest that the human visual system is sensitive to naturally and frequently occurring properties of cast shadows in images (see Kersten, Knill, Mamassian, & Bülthoff, 1996, and Fig. II.20-1b). Recognizing these factors has practical consequences; for example, illumination from above the camera improves the utility of cast shadows in endoscopy (Mishra, Hanna, Brown, & Cuschieri, 2004).

It is well known that, in addition to shadows, the human visual system uses multiple sources of depth information, including stereo disparity, motion parallax, relative brightness, and relative image size to produce estimates of relative depth (Bülthoff & Mallot, 1988; Cutting & Vishton, 1995; Landy, Maloney, Johnston, & Young, 1995). One can use computer graphics to place different sources of information,

or cues, in opposition. For example, Figure II.20-1c shows stereo disparity competing with cast shadows. Which cue wins depends on the weight an individual's visual system gives to the stereo over the shadow information.

In general, the percept experienced is the consequence of how sensory cues are weighted based on their respective reliabilities (Clark & Yuille, 1990; Ernst & Bülthoff, 2004; Jacobs, 1999). This suggests that as we make a visual cue increasingly strong and more diagnostic of a relative depth change (vs. some other cause), we could make visual illusions in which a stronger cue (e.g., the fuzziness of a shadow penumbra) overrides cues that say otherwise. We have already seen this with the green-square demonstration in that the size of the image of the green square does not increase as it approaches the viewpoint, as it normally would under perspective projection.

One way to increase the perceptual strength of shadows for depth is to move the object in the image while keeping the illumination and viewpoint constant. The left top panel of Figure II.20-4a shows the first frame of an animation sequence, and the middle and right panels show two alternative endpoints of the ball's shadow trajectory. While a comparison of the middle and right-hand panels provides some sense of depth differences, the animated versions are considerably more striking (see the Companion Website http://youtube/hdFCJepvJXU). When the shadow travels horizontally in the image, the ball appears to rise above the checkerboard floor, and when the shadow travels diagonally, the ball appears to recede in depth toward the back of the box. The ball's size, appearance, and trajectory are identical in the two animation conditions. The lack of a change in size of the image of the ball presents a cue conflict, which the visual system largely ignores. The bottom row of Figure II.20-4a shows a similar sequence, except that a yellow, solid slab is substituted for the shadow. As long as the slab appears to stay on the floor, viewers typically see the ball rise above the floor for the animation whose shadow endpoint corresponds to the bottom, middle panel but remain on the floor for the endpoint illustrated by the bottom right panel. One explanation for the perception of the different trajectories with the yellow slab is that human vision is sensitive to the correlated motion typical of objects and their shadows resulting from light sources that do not move. This is called the "stationary light source" constraint (Kersten, Mamassian, & Knill, 1997). So even if the slab, ersatz shadow, lacks most of the properties of a shadow, the visual system is equipped to perceive different trajectories as a consequence of the differences in correlated motion. Other factors, such as matching

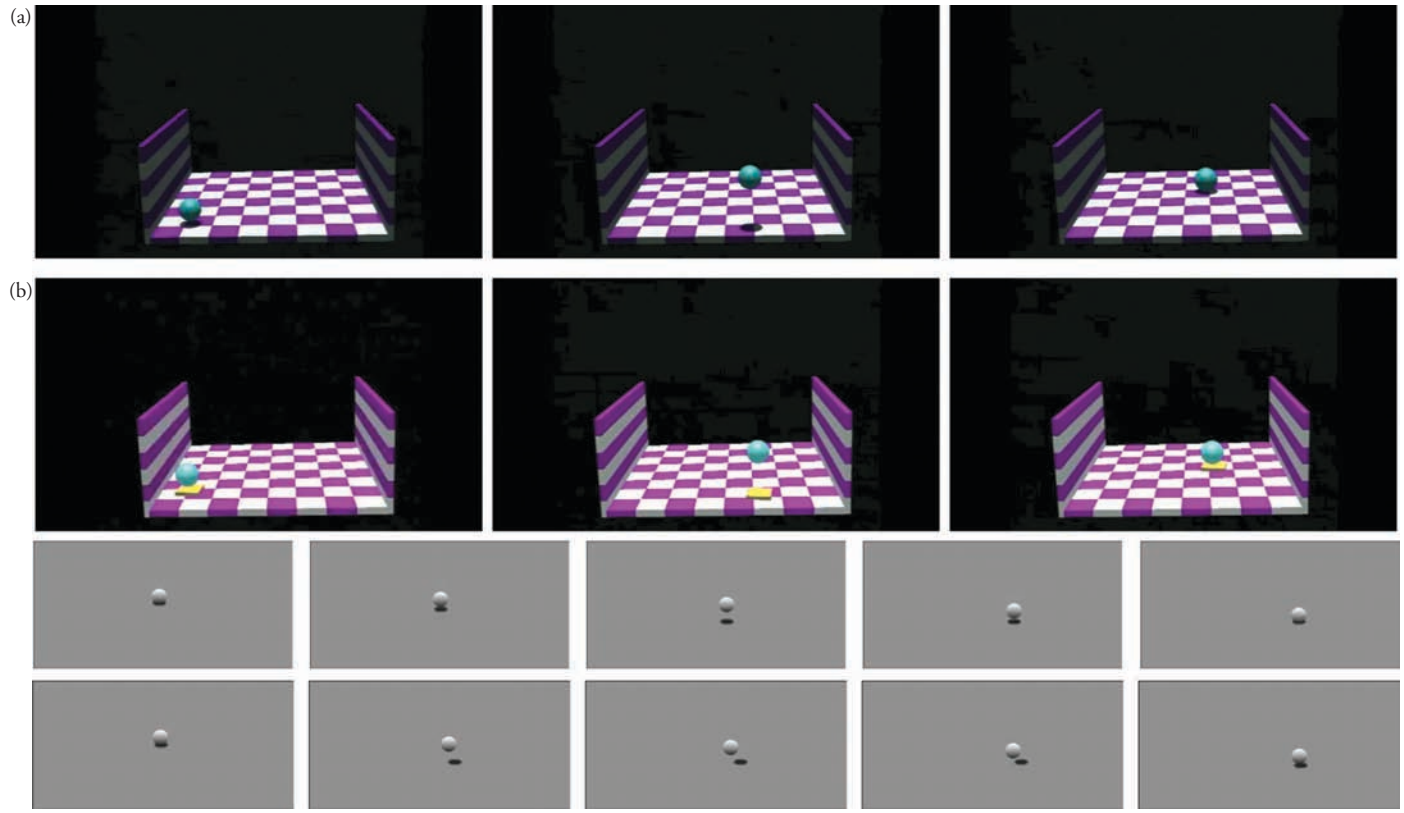


Figure II.20-4. First frame, and the middle and right frames show alternative endpoints of the ball's shadow. The trajectory of the ball is the same in all trajectories. Top panel: When the trajectory of the shadow is horizontal, the ball appears to rise above the floor; when the shadow trajectory is diagonal, the ball appears to travel toward the back of the box. See the Companion Website: http://youtube/hdFCJepvJXU. The perception of the different trajectories in depth relies on the visual system's ability to detect correlated motion (consistent with a stationary light source). Bottom panel: If the shadow is replaced by a yellow slab, the apparent difference in trajectories can still be seen. In addition to correlated motion, treating the slab as a substitute shadow relies on perceiving the slab to be on the floor. The coincidental alignment of one of the primary axes of the slab with that of the floor encourages this. See the Companion Website: http://youtube/TGX7153-Ccw. Figure from Kersten et al. (1997). B: See the Companion Website: http://youtube/6Wbbx6Zi3g4.

sizes between casting object and shadow, also play a role (Ni, Braunstein, & Andersen, 2004) but not as much as the correlated motion.

Another demonstration of the effect of the stationary light source constraint, and its impact on correlated motion, is illustrated in Figure II.20-4b. The top row shows samples of several successive frames in which a ball and its shadow alternately accelerate and decelerate consistent with a fixed light source from above. When all frames are shown in a movie, the ball appears to bounce (see http://youtube/6Wbbx6Zi3g4). However, if the ball has constant velocity (samples shown in bottom row of Fig. II.20-4b), the perceived coupling between the ball and its shadow is lost, the shadow seems to have a life of its own, and the ball no longer appears to bounce.

THE FUTURE OF SHADOWS

Despite their ubiquity, there are substantial computational, perceptual, and neuroscience challenges to our understanding of the visual processing of shadows.

Computational Challenges

While the potential usefulness of cast shadows in computer vision was recognized many years ago (Waltz, 1972), the application to depth estimation has largely been neglected. This stands in contrast to substantial progress in understanding and using other cues such as stereo disparity and motion parallax. To use cast shadows, a visual system must effectively segment a scene into objects and shadows, determine which object (or object part) goes with which shadow (or shadow part; i.e., the shadow "correspondence problem"; Casati, 2008; Mamassian, 2004; Ni et al., 2004), and then use this information in the context of the overall spatial layout (Fig. II.20-5). While there have been a number of analyses of the information provided by shadows (Casati, 2004b; Dee & Santos, 2011; Knill, Mamassian, & Kersten, 1997), the computational challenge will be to turn these observations into useful algorithms.

Statistical constraints on shadow photometry can be used for shadow recognition. For example, the color change across a shadow boundary is largely a change in luminance rather than chromaticity (Finlayson, Drew, & Lu, 2009; Kingdom, Beauce, & Hunter, 2004). Similarly, texture statistics tend to remain similar across shadow boundaries (Zhu, Samuel, Masood, & Tappen, 2010). The global shape of shadow contours have statistical regularities consistent with the shadow patch lying on the surface receiving it. More specifically, shadow shapes follow the geometry of double projection: the shadow on a surface is the projection of the casting object by the light source onto an oriented surface, and the shadow's image is the result of a projection of the surface shadow onto the retina. One consequence of this is that it is not unusual to have a long, thin shadow in an image cast by a compact object (e.g., with the highly oblique sun angles during sunrise or sunset). In contrast, a long, thin image is less likely due to a compact object in the world (Brady & Yuille, 1984).

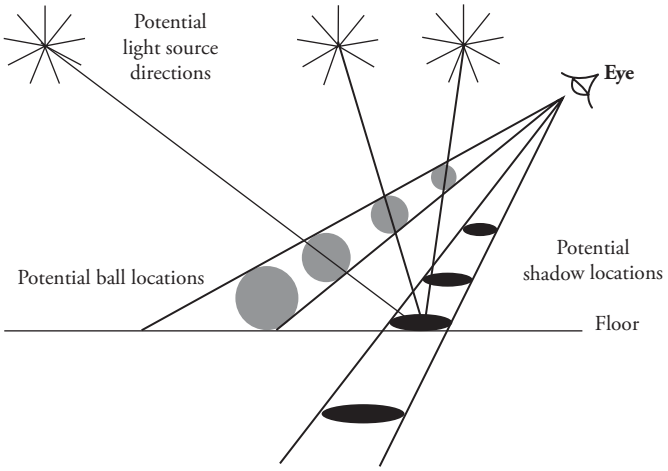


Figure II.20-5. The human visual system solves several problems when estimating depth from cast shadows. Even if the "patches" representing an object and shadow are identified and paired, in order to determine relative depth their locations need to be localized in an appropriate frame of reference (Schrater & Kersten, 2000). For example, while it is obvious that the shadow should be on the ground, an algorithm needs also to determine the ground plane. (From Kersten, 1997.)

Perceptual Challenges

A major problem in visual perception is understanding how the visual system combines multiple cues for a task such as depth perception. One problem in combining shadows with other depth cues is the commensurability of their information. Some cues, like "looming"—the expansion of the size of an object's image on the retina—provide the visual system with information for impending contact between an object and the observer. Rather than depth from the observer, cast shadows more directly provide an estimate of the distance between two objects (the object casting the shadow and the receiving surface; Schrater & Kersten, 2000). In addition, this distance is dependent on usually invisible factors like the light source position. Thus shadows should prima facie not be considered as a metric depth cue like binocular disparity, and this makes it more difficult to understand how these two cues can interact (e.g., Fig. II.20-1c). However, it is conceivable that, given prior assumptions on the statistical structure of the environment, shadows could be "promoted" to a metric depth cue. Such a reliance on the statistical structure of natural scenes has been shown to promote the figure–ground convexity cue (an object with a convex boundary tends to be seen in front) to a full metric depth cue (Burge, Fowlkes, & Banks, 2010).

Neuroscience Challenges

The dual nature of shadows reflects, in large part, two different functional requirements of vision. A basic architectural aspect of the visual brain is the division into two main pathways—the ventral (or temporal) and dorsal (or parietal) streams—believed to be specialized for processing intrinsic object properties and viewer–object spatial relationships, respectively (Milner & Goodale, 2006). As discussed, on the one hand, to recognize objects the visual system

needs to filter out image features, including those, due to shadows, that are not intrinsic to the object. On the other hand, to determine where objects are, we have seen that the visual system uses cast shadows in addition to other depth cues. Bonfiglioli, Pavani, and Castiello (2004) argue that this functional distinction underlies their finding that incongruent shadows had a significant impact on motor task but not on recognition performance. However, it has also been shown that cast shadows can be relatively ineffective in the online control of pointing movements in depth (Hu & Knill, 2011). Consistent with a role for the dorsal stream in processing depth from shadows, neuroimaging results have found that the posterior portion of the medial parietal cortex is activated when observers perceive depth from cast shadows (Katsuyama, Usui, Nose, & Taira, 2011).

Shadows have been studied for millennia, yet there are always new and fascinating observations to be made (Casati, 2012, 2010). Shadow illusions, in particular, raise significant questions about how the brain processes them. Attempting to answer these questions will insure that shadows continue to captivate artists and scientists alike for years to come.

ACKNOWLEDGMENTS

D. K. was supported by the World Class University program funded by the Ministry of Education, Science and Technology through the National Research Foundation of Korea (R31-10008).

NOTES

1. Cast shadows can be distinguished from "attached" or "form" shadows (see Castiello, 2001; Mamassian, Knill, & Kersten, 1998). An attached shadow on an object is the darkened surface that faces away from the light source. A cast shadow on a surface is the dark area created by an object blocking light onto that surface. A cast shadow can also be created when part of an object blocks light onto another part of the same object. For simplicity, in this chapter the word "shadow" usually refers to "cast shadow."
2. The size of the penumbra also depends on the angular size of the light source. On a sunny day, the penumbra caused by the sun is quite small, the sun being far away, and only 1/2° in diameter. On a cloudy day, the effective diameter approaches 180°, and penumbrae are all but invisible except for small depth differences. When there are multiple light sources (e.g., light filtering through the leaves of a tree), there can be multiple shadows that have characteristic transparency-like properties where they overlap.

REFERENCES

Bonfiglioli, C., Pavani, F., & Castiello, U. (2004). Differential effects of cast shadows on perception and action. *Perception*, *33*(11), 1291–1304.

Brady, M., & Yuille, A. (1984). An extremum principle for shape from contour. *Pattern Analysis and Machine Intelligence: IEEE Transactions on Pattern Recognition and Machine Intelligence*, *6*(3), 288–301.

Braje, W. L., Kersten, D., Tarr, M. J., & Troje, N. F. (1998). Illumination effects in face recognition. *Psychobiology*, *26*(4), 371–380.

Braje, W. L., Legge, G. E., & Kersten, D. (2000). Invariant recognition of natural objects in the presence of shadows. *Perception*, *29*(4), 383–398.

Bülthoff, H. H., & Mallot, H. A. (1988). Integration of depth modules: Stereo and shading. *Journal of the Optical Society of America A*, *5*(10), 1749–1758.

Burge, J., Fowlkes, C. C., & Banks, M. S. (2010). Natural-scene statistics predict how the figure–ground cue of convexity affects human depth perception. *Journal of Neuroscience*, *30*(21), 7269–7280.

Casati, R. (2004a). *Shadows: Unlocking their secrets, from Plato to our time*. New York: Vintage.

Casati, R. (2004b). The shadow knows: A primer on the informational structure of cast shadows. *Perception*, *33*(1), 1385–1396.

Casati, R. (2008). The copycat solution to the shadow correspondence problem. *Perception*, *37*(4), 495–503.

Casati, R. (2010). Impossible cast shadows in Ukyio-e paintings. *CARLS Series of Advanced Study of Logic and Sensibility*, *3*, 265–268.

Casati, R. (2012). Some varieties of shadow illusions: Split shadows, occluded shadows, stolen shadows, and shadows of shadows. *Perception*, *41*(3), 357–360.

Castiello, U. (2001). Implicit processing of shadows. *Vision Research*, *41*(18), 2305–2309.

Cavanagh, P. (1991). What's up in top-down processing? In A. Gorea (Ed.), *Representations of vision: Trends and tacit assumptions in vision research* (pp. 295–304). New York: Cambridge University Press.

Cavanagh, P. (2005). The artist as neuroscientist. *Nature*, *434*(7031), 301–307.

Clark, J. J., & Yuille, A. L. (1990). *Data fusion for sensory information processing systems*. New York: Springer.

Cutting, J. E., & Vishton, P. M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In W. Epstein & S. Rogers (Eds.), *Handbook of perception and cognition* (pp. 69–117). San Diego, CA: Academic Press.

Dee, H. M., & Santos, P. E. (2011). The perception and content of cast shadows: An interdisciplinary review. *Spatial Cognition & Computation*, *11*(3), 226–253.

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, *8*(4), 162–169.

Farid, H. (2009). The Lee Harvey Oswald backyard photos: Real or fake? *Perception*, *38*(11), 1731–1734.

Finlayson, G. D., Drew, M. S., & Lu, C. (2009). Entropy minimization for shadow removal. *International Journal of Computer Vision*, *85*(1), 35–57.

Finlayson, G. D., Hordley, S., & Drew, M. (2006). Removing shadows from images. In A. Heyden (Ed.), *Computer Vision—7th European Conference on Computer Vision, 2002* (pp. 129–132). Lecture Notes in Computer Science 2353. Berlin: Springer.

Fiorani, F. (2008). The colors of Leonardo's shadows. *Leonardo*, *41*(3), 271–278.

Gombrich, E. E. H. (1995). *Shadows: The depiction of cast shadows in Western art*. London: National Gallery.

Hu, B., & Knill, D. C. (2011). Binocular and monocular depth cues in online feedback control of 3D pointing movement. *Journal of Vision*, *11*(7): 23, 1–13.

Imura, T., & Tomonaga, M. (2009). Moving shadows contribute to the corridor illusion in a chimpanzee (Pan troglodytes). *Journal of Comparative Psychology*, *123*(3), 280–286.

Imura, T., Yamaguchi, M. K., Kanazawa, S., Shirai, N., Otsuka, Y., Tomonaga, M., & Yagi A. (2006). Perception of motion trajectory of object from the moving cast shadow in infants. *Vision Research*, *46*(5), 652–657.

Jacobs, R. (1999). Optimal integration of texture and motion cues to depth. *Vision Research*, *39*(21), 3621–3629.

Jacobson, J., & Werner, S. (2004). Why cast shadows are expendable: Insensitivity of human observers and the inherent ambiguity of cast shadows in pictorial art. *Perception*, *33*(11), 1369–1383.

Katsuyama, N., Usui, N., Nose, I., & Taira, M. (2011). Perception of object motion in three-dimensional space induced by cast shadows. *NeuroImage*, *54*(1), 485–494.

Kersten, D. (1997). Perceptual categories for spatial layout. *Philosophical Transactions of the Royal Society of London, B*, *352*(1358), 1155–1163.

Kersten, D., Knill, D. C., Mamassian, P., & Bülthoff, I. (1996). Illusory motion from shadows. *Nature*, *379*(6560), 31.

Kersten, D., Mamassian, P., & Knill, D. C. (1997). Moving cast shadows induce apparent motion in depth. *Perception*, *26*(2), 171–192.

Khuu, S., Khambiye, S., & Phu, J. (2012). Detecting the structural form of cast shadows patterns. *Journal of Vision*, *12*(11), 1–10.

Kingdom, F. A. A., Beauce, C., & Hunter, L. (2004). Colour vision brings clarity to shadows. *Perception*, *33*(8), 907–914.

Knill, D. C., Mamassian, P., & Kersten, D. (1997). Geometry of shadows. *Journal of the Optical Society of America A*, *14*(12), 3216–3232.

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research*, *35*(3), 389–412.

Lovell, P. G., Gilchrist, I. D., Tolhurst, D. J., & Troscianko, T. (2009). Search for gross illumination discrepancies in images of natural objects. *Journal of Vision*, *9*(1), 1–14.

Madison, C., Thompson, W., Kersten, D., Shirley, P., & Smits, B. (2001). Use of interreflection and shadow for surface contact. *Perception & Psychophysics*, *63*(2), 187–194.

Mamassian, P. (2004). Impossible shadows and the shadow correspondence problem. *Perception*, *33*(11), 1279–1290.

Mamassian, P., Knill, D. C., & Kersten, D. (1998). The perception of cast shadows. *Trends in Cognitive Sciences*, *2*(8), 288–295.

Milner, D., & Goodale, M. (2006). *The visual brain in action* (2d ed.). Oxford Psychology Series. New York: Oxford University Press.

Mishra, R. K., Hanna, G. B., Brown, S. I., & Cuschieri, A. (2004). Optimum shadow-casting illumination for endoscopic task performance. *Archives of Surgery, 139*(8), 889–892.

Ni, R., Braunstein, M. L., & Andersen, G. J. (2004). Perception of scene layout from optical contact, shadows, and motion. *Perception*, *33*(11), 1305–1318.

Ostrovsky, Y., Cavanagh, P., & Sinha, P. (2005). Perceiving illumination inconsistencies in scenes. *Perception*, *34*(11), 1301–1314.

Porter, G., Tales, A., & Leonards, U. (2010). What makes cast shadows hard to see? *Journal of Vision*, *10*(3), 1–18.

Puerta, A. M. (1989). The power of shadows: Shadow stereopsis. *Journal of the Optical Society of America A*, *6*(2), 309–311.

Rensink, R. A., & Cavanagh, P. (2004). The influence of cast shadows on visual search. *Perception*, *33*(11), 1339–1358.

Schrater, P., & Kersten, D. (2000). How optimal depth cue integration depends on the task. *International Journal of Computer Vision*, *40*(1), 71–89.

Tarr, M., Kersten, D., & Bülthoff, H. H. (1998). Why the visual recognition system might encode the effects of illumination. *Vision Research*, *38*(15–16), 2259–2275.

Taya, S., & Miura, K. (2010). Cast shadow can modulate the judged final position of a moving target. *Attention, Perception, & Psychophysics*, *72*(7), 1930–1937.

Thompson, W. B., Shirley, P., Smits, B., Kersten, D. J., & Madison, C. (1998). *Visual glue*. Technical Report UUCS-98-007. Salt Lake City: University of Utah.

Tomonaga, M., & Imura, T. (2010). Pacman in the sky with shadows: The effect of cast shadows on the perceptual completion of occluded figures by chimpanzees and humans. *Behavioral and Brain Functions*, *6*(1), 38.

Waltz, D. L. (1972). *Generating semantic descriptions from drawings of scenes with shadows.* Technical Report 271. Cambridge, MA: Massachusetts Institute of Technology.

Yonas, A., & Granrud, C. E. (2006). Infants' perception of depth from cast shadows. *Attention, Perception, & Psychophysics*, *68*(1), 154–160.

Zhu, J., Samuel, K. G., Masood, S. Z., & Tappen, M. F. (2010). Learning to recognize shadows in monochromatic natural images. In *Computer Vision and Pattern Recognition: 2010 IEEE Conference* (pp. 223–230). New York: Institute of Electrical and Electronics Engineers.

Chapter 21
The Leaning Tower Illusion

Frederick A. A. Kingdom, Ali Yoonessi, and Elena Gheorghiu

INTRODUCTION

Two identical side by side images of the Leaning Tower of Pisa, photographed from below, appear to rise at different angles (Kingdom, Yoonessi, & Gheorghiu, 2007a,b; see Fig. II.21-1). Although the name of the illusion is a play on the name of the famous building, the illusion works for any image of a receding object. The illusion was discovered by accident in 2007 during a psychophysics (behavioral) experiment conducted by two of the authors. Test observers were shown color-manipulated pairs of images of the same scene and noticed that the pair of Pisa towers appeared different in form not just color. So striking was the effect that it was assumed that the two images had been taken from different photographs. The illusion won an international visual illusion competition in the same year (Kingdom, Yoonessi, & Gheorghiu, 2007c) and has subsequently been featured in numerous books, journals, talks, newspaper articles, calendars, scientific websites, and blogs. Part of the attraction of the leaning tower illusion is its simplicity—one only has to place two identical copies of a photograph of any receding object (e.g., buildings photographed from below, railway lines etc.) side by side to obtain the illusion.

BACKGROUND—THE PERSPECTIVE VIEW

The projection of a three-dimensional scene onto a flat surface such as the retina creates a "perspective view." In the perspective view, parallel contours that recede from the viewer tend to converge toward a common "vanishing" point. Our brains have learned to compensate for these perspective distortions in the retinal image with the result that we tend to perceive objects veridically—that is, correctly. This remarkable ability is not only effortless but unconscious, an example of what the famous 19th-century physicist and physiologist Helmholtz (1866/1962) referred to as "unconscious inference" (pp. 264–301). So good are we at compensating for the perspective

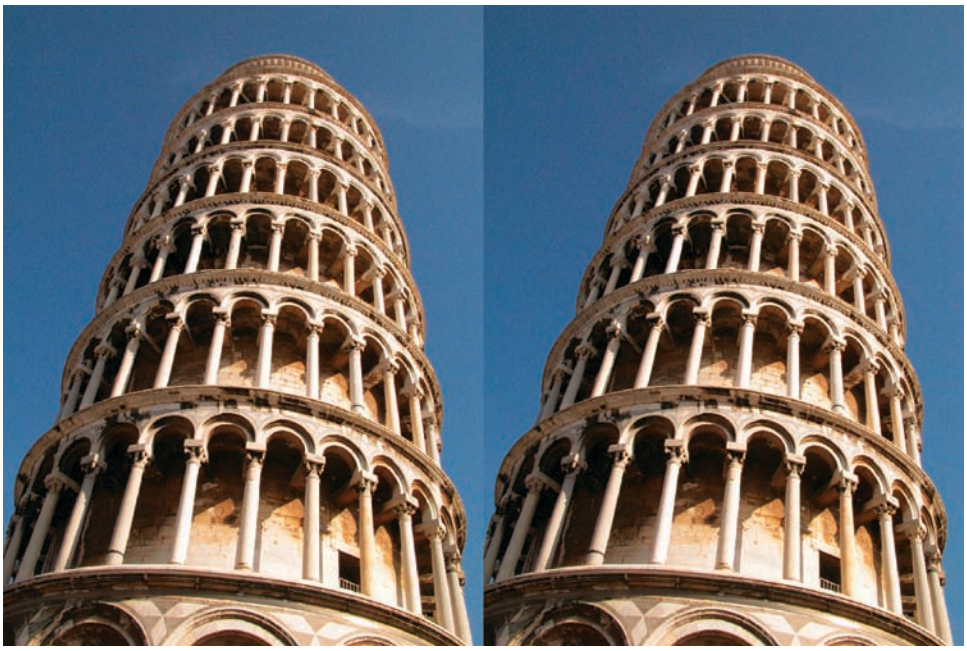


Figure II.21-1. The leaning tower illusion. The two images are identical, yet the one on the right appears more lopsided. The image can be downloaded from the McGill Color Calibrated Image database: http://tabby.vision.mcgill.ca. Original photograph taken by Adriana Olmos.

view, and so unaware are we of doing it, that it is understandable why it was not until the Renaissance, many thousands of years after humans began to draw, that artists worked out how to accurately represent perspective in drawings.

Our brains compensate for perspective not only in the retinal image but in any two-dimensional depiction of a three-dimensional scene, such as a painting, photograph, or movie. Our perception of these depictions, however, is paradoxical: we perceive the scene veridically yet are able to discern the distortions due to perspective. Thus with the Pisa towers, we experience an illusion yet at the same time perceive the perspective distortions that give rise to it.

EXPLANATION OF ILLUSION

Consider the photograph of the Petronas twin towers in Quala Lumpur shown in Figure II.21-2a. The two towers



Figure II.21-2. (a) Photograph of the Petronas twin towers in Quala Lumpur. Although the towers converge toward a common point above the image, we perceive them as rising in parallel. (b) The right-hand tower has been replaced by a copy of the one on the left, so the towers are now parallel on the page, yet appear to diverge. (Photograph courtesy of Thomas Hantner.)

converge in the photograph due to perspective, but our brains compensate for the perspective and as a result we perceive the towers rising in parallel, as they do. It follows that if one views an image in which the two Petronas towers do *not* converge but are parallel, the only reasonable interpretation is that the towers must in reality be *diverging*. We see this in Figure II.21-2b, where the right-hand tower has been replaced by a copy of the tower on the left, such that the two towers are now parallel on the page. The towers appear to diverge. The same principle applies to the Pisa towers.

SIGNIFICANCE OF ILLUSION

If the brain's compensation for perspective was applied separately to each of the Pisa towers in Figure II.21-1, the towers would appear to rise at the same angle. The leaning tower illusion thus reveals that the brain is obliged to compensate for perspective across the whole image rather than separately to its components, even though we know that the components are not part of a single scene. It is as if it is impossible to see the two Pisa towers as separate, albeit identical, versions of the same image: instead one sees the "Twin Towers of Pisa," with the "correct" interpretation that one tower leans more than the other.

Although the leaning tower illusion is a misperception, it does not imply that our brains are flawed. Our normal visual experience is of a three-dimensional world that fills our field of view. It would seem sensible for the visual system to apply perspective correction to the whole visual scene rather than separately to its parts. Otherwise, when deriving a coherent and accurate representation of the scene, conflicting interpretations might arise that would be difficult to reconcile.

OTHER EXPLANATIONS FOR ILLUSION

Maniatis (2008) has suggested that the leaning tower illusion is not primarily caused by perspective but by a process akin to acute-angle expansion. Acute-angle expansion is the phenomenon in which acute angles tend to appear larger than they are (Kitaoka & Ishihara, 2000). If the angle subtended by the inner two adjacent sides of the towers in Figure II.21-1 were to be expanded perceptually, this could result in the impression that the towers were diverging. Akiyoshi Kitaoka (2007) was the first to point out this possibility and designed Figure II.21-3a to test it. The arrangement of the angles in the cartoon figures is similar to that in the Pisa towers, but the figures do not appear to recede. If the illusion was due to acute-angle expansion one should experience a similar-sized illusion with the cartoon figures, whereas if the perspective explanation is correct, one should not see the illusion. The illusion is tiny or nonexistent in the cartoon figures, thereby supporting the perspective explanation.

Figure II.21-3b, shows that the illusion is not an intrinsic property of the Pisa tower: the illusion is not observed when the tower is photographed from a sufficient distance such that it does not appear to be a receding object. Johnson (2011) has provided additional evidence for the perspective explanation by showing that the strength of the illusion can be altered by manipulating its perspective cues.

RELATIONSHIP TO SIZE-CONSTANCY ILLUSIONS

An important class of visual illusion pertains to the mechanisms for achieving size constancy. Size constancy refers to the observation that the perceived size of an object remains constant with distance in spite of changes in retinal-image size. As a consequence, if one places an object of constant retinal-image size at various apparent distances in a scene, it will appear to differ in size, as exemplified by the Ponzo illusion in Figure II.21-4.

Figure II.21-5 illustrates the size-constancy illusion with an image of an elongated cube. The cross-sectional diameter of the cube on the page is the same at both ends (readers can check this with a ruler), yet the cube appears to increase in diameter with distance. The figure is also an illusion of perspective. The elongated sides of the cube are parallel on the page as the cube recedes, yet the sides appear to diverge as with the corresponding sides of the two Pisa towers in Figure II.21-1. Is the leaning tower illusion therefore a size-constancy illusion? If by "size" one means the distance between corresponding sides of the two Pisa towers, then the answer is yes, because that distance remains constant on the page as the towers recede. However, an equivalence between size relations and perspective geometry does not imply that perspective and size-constancy illusions are mediated by the same mechanism. Palmer (1999) for example has suggested that the perceived size of an object is determined by its retinal-image size relative to that of its neighbors in the same depth plane, whereas perspective is determined by the angles subtended by converging lines. In Figure II.21-5, however, there are no neighboring objects to act as references for the computation of relative size, so the figure tends to support the view that there is an intimate relationship between illusions of size-constancy and illusions of perspective.

LIMITS TO THE ILLUSION

If a series of photographs of the Pisa tower are placed in a row, one might suppose that the illusion would accumulate, leading to the comical prediction that the tower on the far right should topple over altogether! Figure II.21-6, however, shows that this does not happen. The illusion can be observed in any pair of attended images, but there is no impression that the towers become increasingly lopsided from left to right. Why not? For the illusion to accumulate, the perspective-correction would have to be applied not to the raw image of each tower but instead to its already-illusory internal representation, which is arguably an implausible scenario.

OTHER EXAMPLES OF ILLUSION

The illusion works with any image of a receding object, such as the identical pair of images of tram lines shown in Figure II.21-7.

(a)



(b)

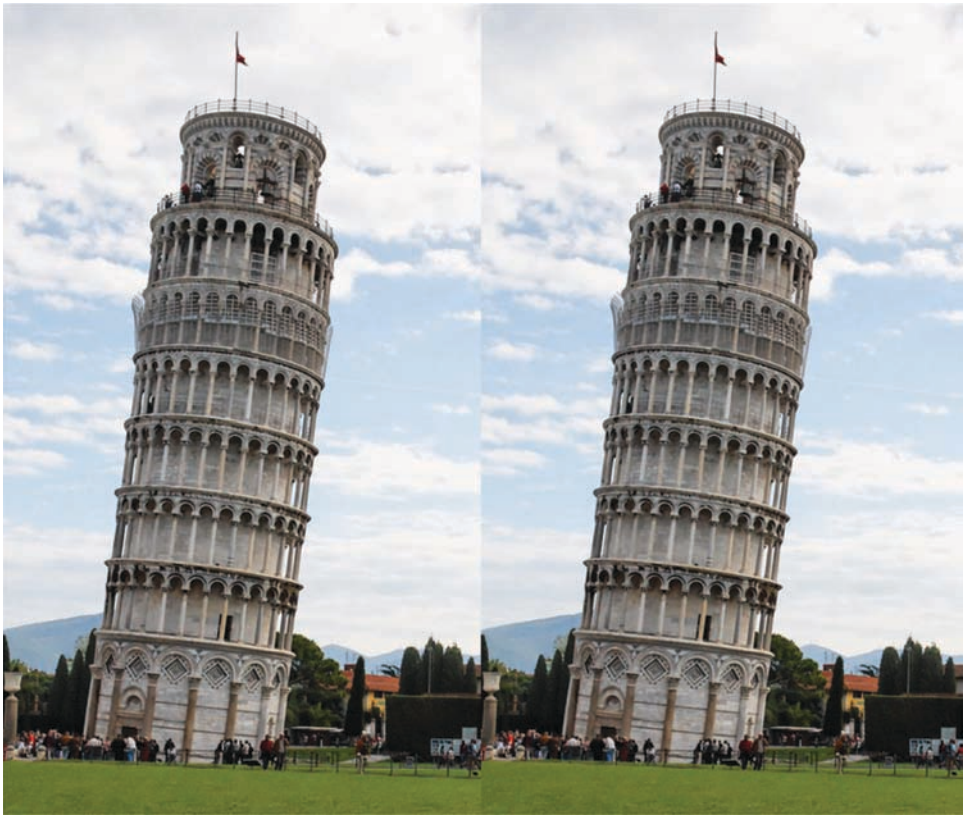


Figure II.21-3. (a) The outlines of two identical cartoon figures describe similar angles to those of the Pisa towers in Figure II.21-1, yet no illusion obtains. Unlike the Pisa towers in Figure II.21-1, the cartoon figures do not appear to recede. (Figure courtesy of Akiyoshi Kitaoka.) (b) The illusion is also absent in two identical images of Pisa towers that do not appear to recede, ruling out that the illusion is caused by the particular structure of the towers. (Figure courtesy of Dejan Todorovic.)

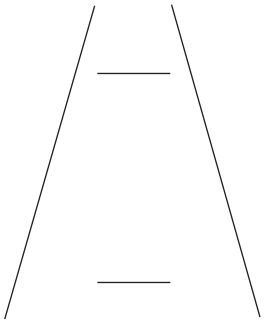


Figure II.21-4. Ponzo illusion. The two horizontal lines are identical in length yet appear to have different lengths, because they appear to lie at different depths from the observer.

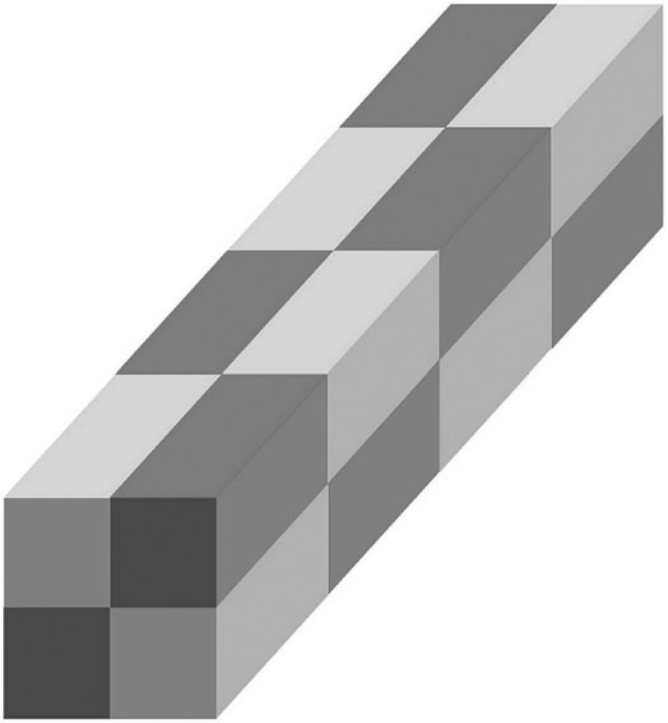


Figure II.21-5. Elongated cube illusion. The elongated edges of the cube run parallel on the page yet appear to diverge with distance—an illusion of perspective closely related to the leaning tower illusion. The horizontal width of the cube appears to increase with distance even though constant, an illusion of size constancy. Thus the elongated cube shows the intimate relationship between illusions of perspective and of size.

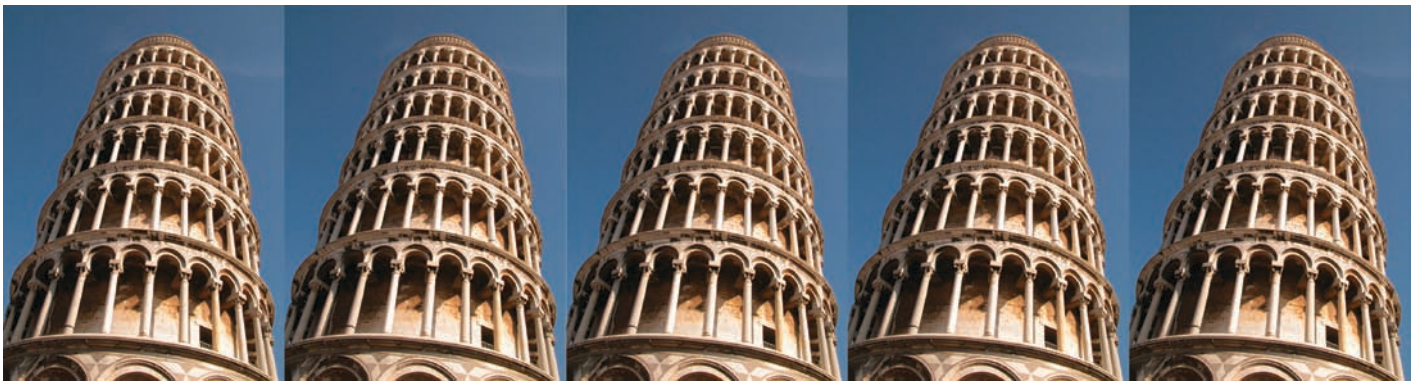


Figure II.21-6. When multiple copies of the Pisa tower are placed next to one another, the illusion can still be seen in any pair of towers but does not accumulate across the screen.



Figure II.21-7. Identical images of tram lines that appear to run in different directions. (Photo courtesy of Brian Micklethwait.)

REFERENCES

Helmholtz, H. von. (1962). *Treatise on Physiological Optics. Vol II* (J. P. L. Southall, Trans.). New York: Dover. (Original work published 1866)

Johnson, A. (2011). Varying image perspective weakens the leaning tower illusion. *Journal of Vision*, *11*(11), 1116 (abstract).

Kingdom, F. A. A., Yoonessi, A., & Gheorghiu, E. (2007a). Leaning tower illusion. *Scholarpedia*. http://www.scholarpedia.org/article/Leaning_tower_illusion

Kingdom, F. A. A., Yoonessi, A. & Gheorghiu, E. (2007b). The leaning tower illusion: A new illusion of perspective. *Perception*, *36*, 475–477.

Kingdom, F. A. A., Yoonessi, A. & Gheorghiu, E. (2007c). The leaning tower illusion. Winner: Best illusion of the year contest. http://illusionoftheyear.com/2007/the-leaning-tower-illusion/

Kitaoka, A. (2007). Illusion news. (http://www.psy.ritsumei.ac.jp/~akitaoka/illnews5e.html

Kitaoka, A., & Ishihara, M. (2000). Three elemental illusions determine the Zollner illusion. *Perception & Psychophysics*, *62*, 569–575.

Maniatis, L. M. (2008). The leaning tower illusion is not a simple perspective illusion. *Perception*, *37*, 1769–1772.

Palmer, S. E. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA: MIT Press.

Yoonessi, A., Gheorghiu, E., & Kingdom, F. A. A. (2008). The leaning tower illusion: Orientation contrast or perspective distortion? Reply to Maniatis." *Perception*, *37*, 1773–1775.

Chapter 22
The Invisible Saddle, or the Cap-or-Cup Illusion

Jan Koenderink, Andrea van Doorn, and Johan Wagemans

THE CAP-OR-CUP ILLUSION

Consider Figure II.22-1 (left). It is a shaded view of the surface patch shown at Figure II.22-1 (right), although this might be hard to believe.

The surface is assumed Lambertian (roughly white plaster) and as illuminated by a uniform, collimated beam (roughly direct sunlight). This surface shape is conventionally referred to as a "saddle." To most observers the image in Figure II.22-1 (left) appears as a cap (spherical convexity), sometimes as a cup (spherical concavity), but never as a saddle (we collected thousands of reports; see Wagemans, van Doorn, & Koenderink, 2010). Apparently the immediate visual awareness is illusory because it is highly nonveridical. This is the case of "the invisible saddle."

It is possible to illuminate a cap in such a way that it would yield an image identical to Figure II.22-1 (left). It is also possible to do this with a cup; in fact, one could illuminate any quadric surface (Fig. II.22-2) in such a way that it would yield this same image (Koenderink & van Doorn, 2003b). The image has infinitely many interpretations that are physically possible (Belhumeur, Kriegman, & Yuille, 1999; Forsyth & Ponce, 2002; Horn, 1970). Thus a cap-or-cup interpretation has probability zero to apply. Yet these are the only reports one obtains. Thus the invisible saddle might also be called the "cap-or-cup illusion."

The cap-or-cup illusion is so strong that even vision scientists have disregarded the saddle interpretation since the dawn of their field (Brewster, 1832; Hess, 1950; Hershberger, 1970; Palmer, 1999), thus entirely misrepresenting the response repertoire (Ramachandran, 1988a, 1988b). Only recently has this been rectified (Wagemans et al., 2010). The cap-or-cup illusion is a strong one, since it does not go away even if we *know* we are looking at a saddle. (Fig. II.22-1, left, is a veridical rendering of a saddle; just try to *see* it.) Apparently it arises in precognitive awareness and is "cognitively impenetrable" (Fodor, 1983). In that respect, it is a truly *visual* illusion. In this chapter we trace some of the possible causes and conceptual consequences of the cap-or-cup illusion.

ECOLOGICAL CONSIDERATIONS

Might it be the case that saddles are extremely sparse in the real world? This is the argument of ecological validity (Gibson, 1966). A Bayesian argument (Bayes, 1763; Knill & Richards, 1996) would at least go some way to account for the cap-or-cup illusion. Saddles are indeed special in that a smooth object, of finite size, cannot have a completely saddle-like surface, like a sphere that is completely made up of caps (Hilbert, 1901).

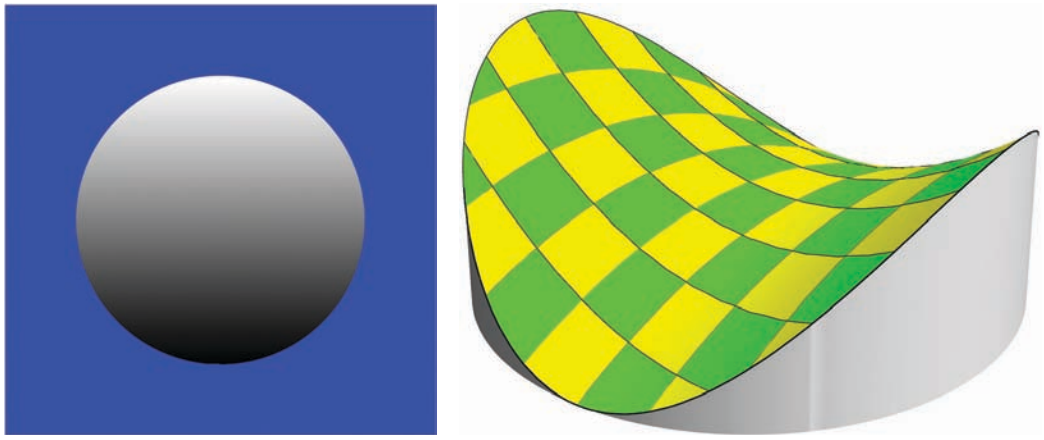


Figure II.22-1. Left: Picture of an illuminated surface seen against the blue sky. Right: 3D impression of the surface: it is a saddle shape.

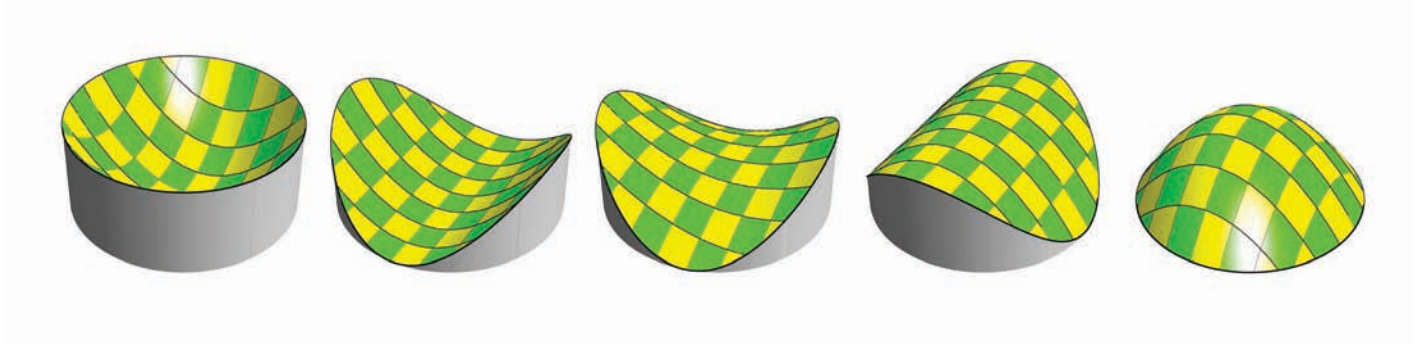


Figure II.22-2. The zoo of quadric shapes (Koenderink, 1990; Koenderink & van Doorn, 1992). Notice that any surface looks like one of the quadric shapes in a small enough environment of any of its points: this is a complete catalogue except for the fact that it is continuous. We must interpolate shapes between the two endpoints of the scale; here we give only some intermediate landmarks. These are (from left to right), a cup (endpoint), a rut, a symmetrical saddle, a ridge, and a cap (the other endpoint). All shapes between rut and ridge are (typically) asymmetrical saddles, all points between cup and rut are asymmetrical cups, and all points between ridge and cap are asymmetrical caps.

One way to approach this problem is to consider the counts of peaks (caps), pits (cups), and passes (saddles) on a single connected island (such as Australia). From differential topology we see that the number of passes equals the number of peaks plus the number of pits minus two (Guillemin & Pollack, 1974). Maxwell (1870), in a famous paper "On Hills and Dales," shows how to divide a landscape into either hills or dales. He finds that the number of saddles equals the number of hills plus one, or—equivalently—that the number of bars (also saddles) equals the number of dales plus one. This suggests that saddles (or bars) are not rare compared to caps or cups.

We have found (Koenderink & van Doorn, 2003a; Lillholm & Griffin, 2009; Fig. II.22-3) the probability density function of any shape on a random Gaussian surface (Fig. II.22-4). It is a bimodal distribution that peaks at the ridges and ruts (convex and concave cylinders), is zero at the umbilicals (caps or cups), and has a finite value at the symmetrical saddles. We find that the larger area of the surface is covered with saddles. Thus the cap-or-cup illusion cannot be "explained" on the basis of ecological considerations.

INVERSE OPTICS

Can "inverse optics" account (Poggio, 1984) for the cap-or-cup illusion? It certainly explains why all quadric shapes are in a single metameric class when the direction of illumination is unknown (Koenderink and van Doorn, 2003b). It also yields a relation between the direction of illumination and the veridical interpretation given the usual constraints (Lambertian surface, collimated, uniform illumination, frontal view; Horn, 1970; Lambert, 1760). In the case of cups, the direction of illumination is in the direction of the shading gradient; in the case of caps, it is in the opposite direction; and in the case of a symmetrical saddle, it is at right angles to the gradient. In the case of cylinders, it is undefined, with the shading gradient orthogonal to the cylinder axes, independent of the direction of illumination.

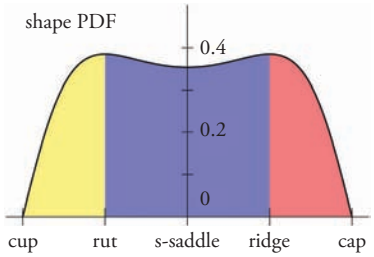


Figure II.22-3. The probability density function (PDF) for the quadric shapes on a random Gaussian surface (Fig. II.22-4). The symmetrical saddle ("s-saddle") is at the center. The yellow area indicates (asymmetrical) cups, the red area (asymmetrical) caps, and the blue area (asymmetrical) saddles. Integrating over the blue area we find that 57.8% of the total surface area is covered with saddles.

Thus the direction of illumination cannot be inferred from a picture like Figure II.22-1 (left). It can be inferred (in a statistical sense) from the image of a sufficiently articulated surface, like a Gaussian random surface (Koenderink & Pont, 2003; Koenderink, van Doorn, & Pont, 2004). In such cases it might indeed be held that one "sees saddles" because one sees the articulated surface, which is composed of a continuous distribution of quadric shapes. This is the conventional interpretation. However, an objection might be that one perhaps sees cups and caps, and some kind of "glue" that binds them together.

There is some anecdotal evidence in favor of this interpretation (which may appear weird at first blush), because sculptors tend to work on the convexities and concavities, letting the saddle-like parts care for themselves (Rogers, 1969). This might be objectified by considering copies of sculpture and measuring variations in the convex, concave, and saddle-like parts. Moreover, in many cultures the saddle-like "glue" is simply left out, and the surface becomes a conglomerate of caps, with sharp V grooves between them.

THE SHADING CUE

Given the basic shape metamerism revealed by inverse optics analysis (Belhumeur et al., 1999; Koenderink & van Doorn, 2003b), one expects the "shading cue" to be a relatively weak one. This is apparently contradicted by the fact that the large majority of observers are visually aware of a cap or cup when looking at the picture shown in Figure II.22-1 (left; Hershberger, 1970; Hess, 1950; Ramachandran, 1988a, 1988b; Wagemans et al., 2010).

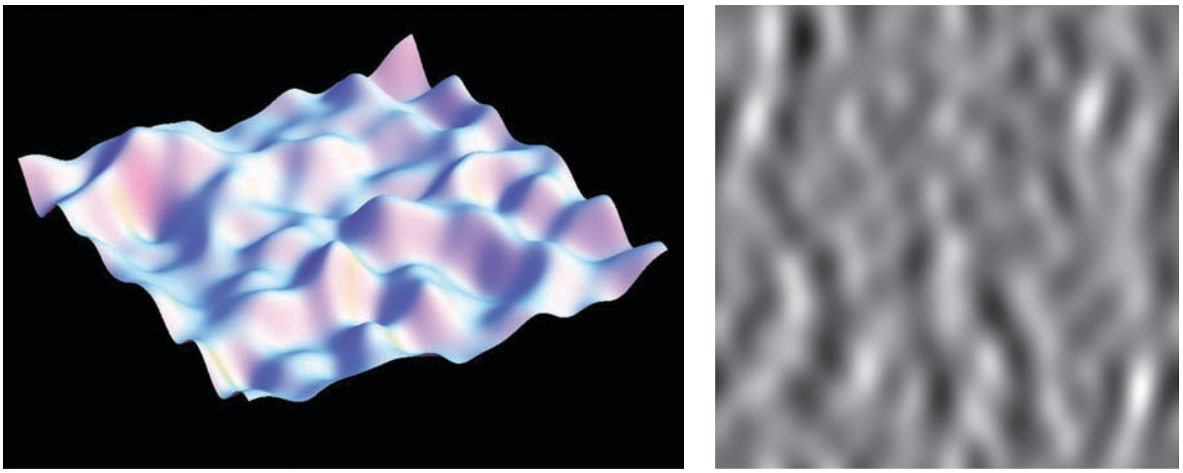


Figure II.22-4. Left: 3D impression of a random Gaussian surface. Right: An image of the surface (considered Lambertian) illuminated by a collimated, uniform beam. There exist obvious pits (cups) and peaks (caps). Does one "see" saddles or merely "glue" between the caps and cups?

Only a minority of observers report experiencing it as a flat distribution of pigments on the page, a veridical perception if there ever was one! (See Wagemans et al., 2010.)

That this finding might carry little weight becomes evident as one tries to isolate the shading cue proper. The picture (Fig. II.22-1, left) is conventionally presented as a case of "pure shading," but it is actually more complicated. For instance, the circular boundary is often (precognitively) interpreted as flat and frontoparallel, thus forcing cap or cup.

This is in itself somewhat interesting, because in Figure II.22-1 (left) we intentionally show the object against a "blue sky." This rules out that the outline is an edge between the shape and a flat pedestal (a uniform gray background surely indicates a flat pedestal; Fig. II.22-5, right). This leaves an occluding contour or a "flag edge" as possibilities. Either of these can be a general (twisted) space curve. Perhaps this is not seen because human observers generally (but erroneously) tend to interpret occluding contours as planar. (Even David Marr [1982] committed this error, apparently not noticing that the planar case is quite nongeneric; Koenderink & van Doorn, 1982.)

One way to "isolate" the shading cue is to blur the outline. One obtains a picture like Figure II.22-6. A formal psychophysical study suggests that shading proper is only a very weak shape cue (Erens, Kappers, & Koenderink, 1993). Apparently most cases of purportedly "shape from shading" are really to be considered "shape from shading augmented by X," where X could be implemented in a variety of ways, as discussed next.

THE AUGMENTED SHADING CUE

If "shape from shading" is really "shape from shading augmented by X," then one might try to influence the cap-or-cup illusion, or perhaps even destroy it, by skillful

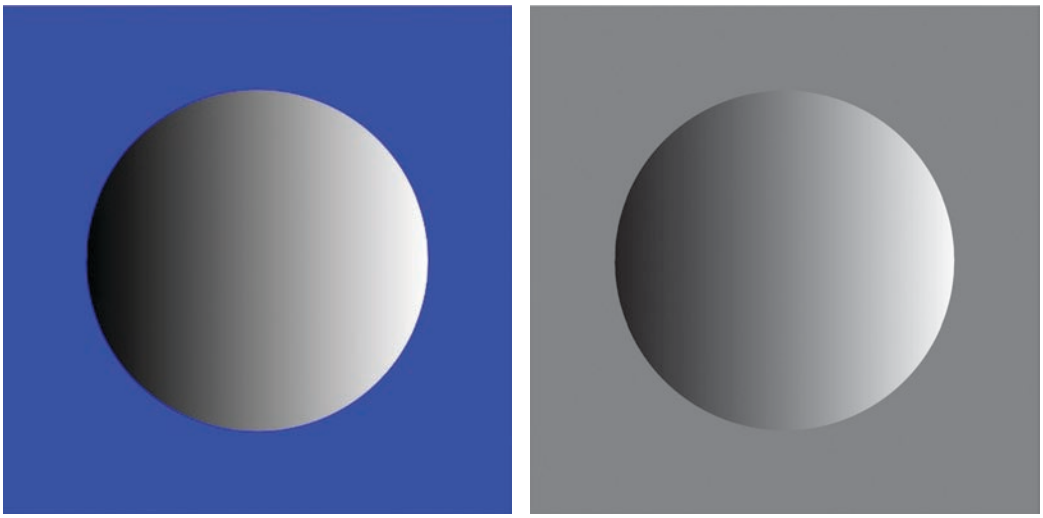


Figure II.22-5. Left: The basic image against a "blue sky." Right: The basic image against a gray background. In the latter case the background is seen as planar (the only solution consistent with the shading) and the plane as forming a unity with the cap or cup. Then the outline must be a planar edge. In the left image the outline cannot be an edge; it must be an occluding contour or a flag edge. In either case, planarity is not implied.

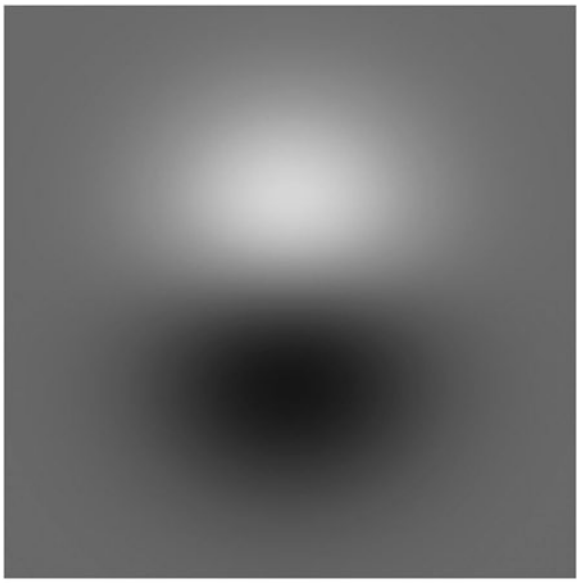


Figure II.22-6. This image is similar to Figure II.22-1, left, but with a blurred outline. Many observers can see it as a Gaussian hill (cap) or valley (cup), although the impression is not very strong. No one sees it as a saddle illuminated horizontally, although that is a perfectly veridical interpretation too.

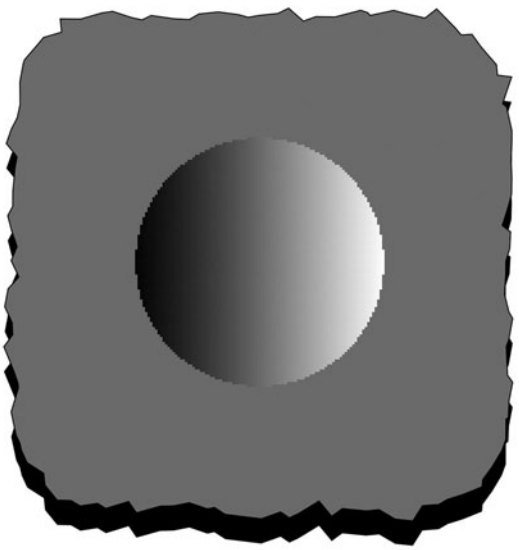


Figure II.22-7. The pedestal casts a shadow on the white of the paper, clearly identifying the direction of illumination as from above. The shading gradient is at right angles to it, thus the shape is a symmetrical saddle. No one sees this though; only caps and cups are reported.

manipulation of the extraneous cue(s). In this section we consider some obvious variations.

One obvious choice for X is to indicate the direction of illumination by some visual means (Fig. II.22-7). It fails to work. Observers prefer to see two mutually disparate directions of illumination.

Another choice for X is to change the outline. This has a strong effect (Fig. II.22-8; Koenderink & van Doorn, 2003b). It is easy enough to turn a cap into a ridge, or a cup into a rut. However, it does not allow one to produce a saddle (we tried numerous combinations).

Changing the outline into a solid rim also has major consequences (Fig. II.22-9). This vetoes the occluding contour interpretation and forces a flag edge. Shading the rim lets one modulate its spatial attitude; this yields an important additional handle on the image. Thus one easily forces seeing a cup over a cap, and this easily overrides the preference for illumination from above, which is anyway weak, or even absent in many observers (Wagemans et al., 2010). However, this does not serve to force the perception of a saddle (Wagemans et al., 2010).

In the example shown in Figure II.22-10 we attempt to summon up what we believe to be the strongest determinant of shape so as to force the shading cue. It does not work. Observers apparently prefer to have nonveridical and even internally inconsistent perceptions over seeing a saddle

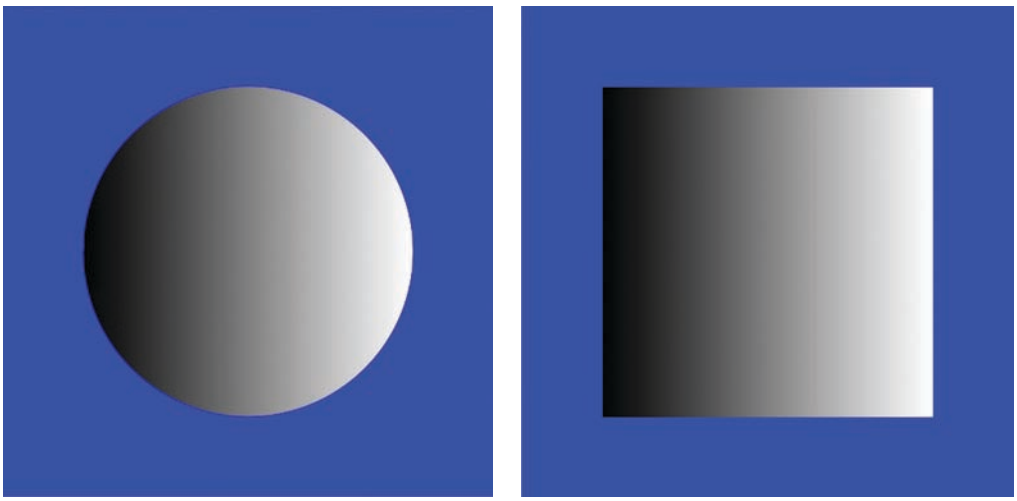


Figure II.22-8. By changing the circular outline (left) into a square outline (right), the shape in immediate visual awareness changes from a cap (sometimes a cup) to a ridge (sometimes a rut). Apparently the outline has a strong influence on the shape one is aware of, for the shading (a linear gradient) is identical in both cases.



Figure II.22-9. The edge quality may determine the cap/cup decision (left). The edge looks like a thick rim and partakes in the shading. In the case of the square outline, it forces the ridge/rut decision (center). It does not seem to work as well to force a saddle impression though.

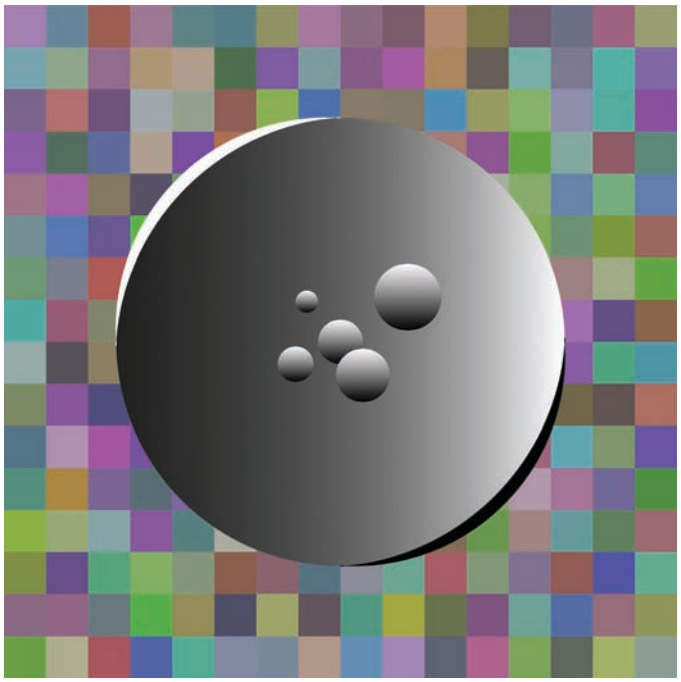


Figure II.22-10. Our most sophisticated shot at depicting a saddle. It evidently fails: saddles appear to be "invisible." Here the light direction is indicated by the spheres (vertical; most observers find it compelling), and the outline is clearly indicated as a flag edge (most observers report a thick, curved tile). The shading gradient is horizontal, thus the shape is a symmetrical saddle, an interpretation that agrees with all the evidence. Yet observers report a rut with cylinder axis along the lower left to upper right corner, at odds with the shading (a rut the shading gradient should be along the upper left to lower right diagonal). Apparently one prefers an inconsistency over seeing a saddle.

shape. The conclusion is that saddles are indeed invisible, at least in immediate (precognitive) visual awareness.

CONCLUSIONS

The most likely explanation of the cap-or-cup illusion is the dominating influence of the circular outline, which is almost invariably, and fully automatically, seen as planar and frontoparallel. This forces the shape to be cap or cup. Notice that there is no geometrical reason for the circle to correspond to a planar, frontoparallel circle. It might just as well be an oblique ellipse, an arbitrary curve on the surface of a right circular cone with apex at the eye, or even a (dense enough) set of mutually disperse points on such a surface (Fig. II.22-11). However, from an ecological perspective the latter interpretations have a probability of zero, namely the probability of the eye to coincide with the apex of the cone. General triaxial ellipsoids can have circular contours (Fig. II.22-12), corresponding to oblique, planar ellipses. However, in order to obtain such a view the eye has to be on a certain curve in space, thus the probability is again zero.

Twiddling with the circular outline should remove the pronounced preference for cups or caps. Indeed it does (Fig. II.22-8). For instance, in the case of a square outline, lined up with the shading gradient, observers report ruts or ridges. Thus this is another variety of the invisible saddle, which may be called the "ridge-or-rut illusion."

In this case the illusion is apparently due to the fact that two edges of the square are taken as mutually parallel and frontoparallel, whereas the two other edges are seen as curved (Fig. II.22-13). In this case the saddle would require one to see the edges as straight, though mutually skew. This is hardly less probable than seeing two edges as straight and parallel and two others curved, it seems. Thus the ridge-or-rut illusion, which is just as compelling as the cap-or-cup illusion, is an even stronger indication for the "invisibility" of saddles.

The cup/cap ambiguity can be lifted (forcing a cup; Fig. II.22-9) by drawing a thick tile instead of a mere flimsy surface, but this does not (or at least not for many observers) work in the case of the saddle-in-a-square, again suggesting that the "invisibility" of saddles is deeply ingrained in the human visual system. For immediate visual awareness, saddle shapes are perhaps better said to not exist. Different from caps, cups, ridges, or ruts, they are not part of the repertoire of possible shapes. This conclusion is fortified by the historical fact that saddles do not appear in the "complete" catalogue of surface shapes drawn up by

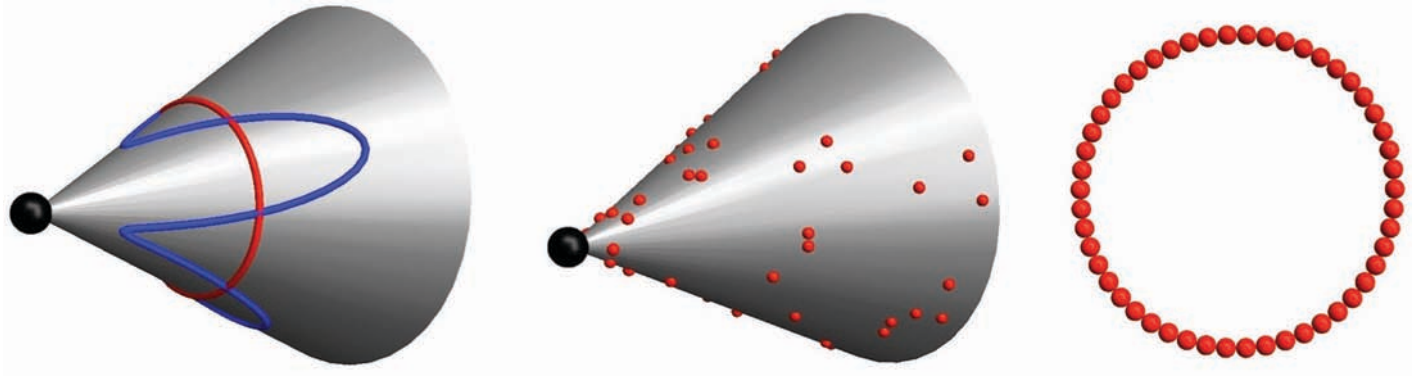


Figure II.22-11. Left: Two conical curves (red and blue) that both look circular from the apex (black). The blue curve is evidently not planar. Right: Two views of the same disperse point set. Seen from the apex the point set (which is not even a curve) appears as a (frontoparallel!) circle.

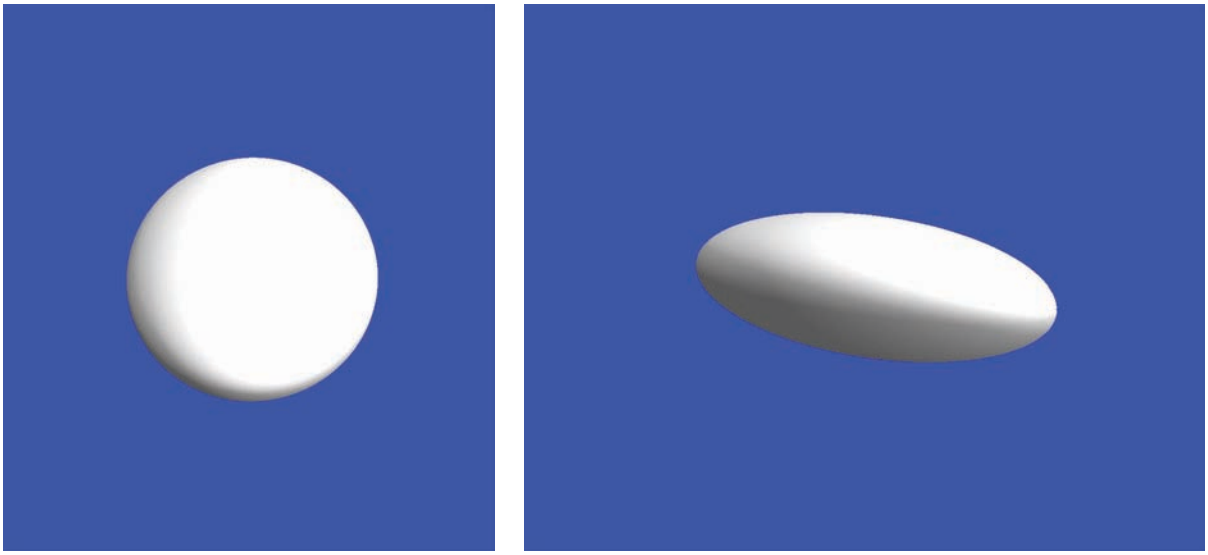


Figure II.22-12. Two views of the same triaxial ellipsoid. In the view at the left, the outline is circular. It is not frontoparallel though.

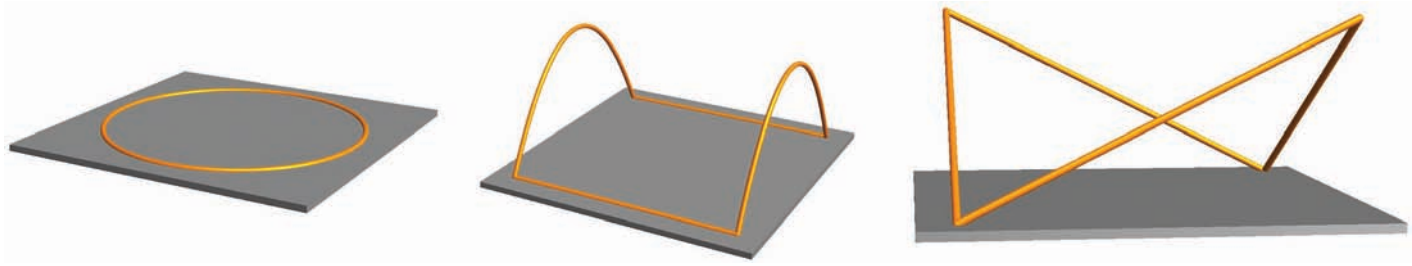


Figure II.22-13. A flat rim; a partwise curved, globally nonplanar rim; and a partwise straight, globally non-planar rim. The left case corresponds to the cap and the center to the ridge impression, whereas the right case would apply to a saddle impression. The latter case, of course, never occurs. Why? Geometrically it is a perfectly viable option.

Leon Battista Alberti (1435) in the 15th century in a work that is often (and in our view rightly) considered "scientific." The saddles were introduced by Carl Friedrich Gauß (1827) in the first full inventory of local surface shapes, which dates from the first half of the 19th century.

The bottom line is this: humans are blind to saddles, and the cap-or-cup illusion is a consequence of that and of the fact that circles look frontoparallel in visual space.

REFERENCES

Alberti, L. B. On Painting. [First appeared 1435–36]. Translated with Introduction and Notes by John R. Spencer. New Haven: Yale University Press. 1970 [First printed 1956].

Bayes, T [communicated by Mr. Price] (1763). An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, *53*, 370–418.

Belhumeur, P., Kriegman, D., & Yuille, A. (1999). The bas-relief ambiguity. *International Journal of Computer Vision*, *35*, 33–44.

Brewster, D. (1832). *Letters on natural magic*. London: John Murray.

Erens, R. G. F., Kappers, A. M. L., & Koenderink, J. J. (1993). Perception of local shape from shading. *Perception & Psychophysics*, *54*, 145–156.

Fodor, J. A. (1983). *Modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.

Forsyth, D. A., & Ponce, J. (2002). *Computer vision: A modern approach*. Upper Saddle River, NJ: Prentice Hall.

Gauß, C. F. (1827). Disquisitiones generates circa superficies curvas. In *Commentationes Societatis Regiae Scientiarum Gottingesis Recentiores* (Vol. VI, pp. 99–146). Gottingae: Typis Dieterichianis. (English translation: J. C. Morehead & A. M. Hiltebeitel (Trans. and Eds.), *General investigations of curved surfaces of 1827 and 1825*. Princeton, NJ: Princeton University Library, 1902.)

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.

Guillemin, V., & Pollack, A. (1974). *Differential topology*. New York: Prentice Hall.

Hershberger, W. A. (1970). Attached shadow orientation perceived as depth by chickens reared in an environment illuminated from below. *Journal of Comparative and Physiological Psychology*, *73*, 407–411.

Hess, E. H. (1950). Development of the chick's responses to light and shade cues of depth. *Journal of Comparative and Physiological Psychology*, *43*, 112–122.

Hilbert, D. (1901). Über Flächen von konstanter Krümmung. *Transactions of the American Mathematical Society*, *2*, 87–99.

Horn, B. K. P. (1970). *Shape from shading: A method for obtaining the shape of a smooth opaque object from one view* (PhD thesis). Massachusetts Institute of Technology, Cambridge, MA.

Knill, D. C., & Richards, W. (1996). *Perception as Bayesian inference*. Cambridge, UK: Cambridge University Press.

Koenderink, J. J. (1990). *Solid shape*. Cambridge, MA: MIT Press.

Koenderink, J. J., & Pont S. C. (2003). Irradiation direction from texture. *Journal of the Optical Society of America A*, *20*, 1875–1882.

Koenderink, J. J., & van Doorn, A. J. (1982). The shape of smooth objects and the way contours end. *Perception*, *11*, 129–137.

Koenderink, J. J., & van Doorn, A. J. (1992). Surface shape and curvature scales. *Image and Vision Computing*, *10*, 557–565.

Koenderink, J. J., & van Doorn, A. J. (2003a). Local structure of Gaussian texture. *IEICE Transactions on Information and Systems*, *86*(7), 1165–1171.

Koenderink, J. J., & van Doorn, A. J. (2003b). Shape and shading. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neurosciences* (pp. 1090–1105). Cambridge, MA: MIT Press.

Koenderink, J. J., van Doorn, A. J., & Pont, S. C. (2004). Light direction from shad(ow)ed random Gaussian surfaces. *Perception*, *33*, 1405–1420.

Lambert, J. H. (1760). *Photometria, sive, De mensura et gradibus luminis, colorum et umbrae* Augustae Vindelicorum, Germany: V. E. Klett.

Lillholm, M., & Griffin, L. D. (2009). Statistics and category systems for the shape index descriptor of local 2nd order natural image structure. *Image and Vision Computing*, *27*, 771–781.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York: Freeman.

Maxwell, J. C. (1870). On hills and dales. *Philos. Mag.* (Maxwell's Collected Works), *40*(269), 421–427.

Palmer, S. E. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA: MIT Press.

Poggio, T. (1984). Low-level vision as inverse optics. In M. Rauk (Ed.), *Proceedings of symposium on computational models of hearing and vision* (pp. 123–127). Tallinn, Estonia: Academy of Sciences of the Estonian S.S.R.

Ramachandran, V. S. (1988a). Perceiving shape from shading. *Scientific American*, *259*, 58–65.

Ramachandran, V. S. (1988b). Perception of shape from shading. *Nature*, *331*, 163–166.

Rogers, L. R. (1969). *Sculpture, the appreciation of the arts*. Vol. 2. Oxford: Oxford University Press.

Wagemans, J., van Doorn, A. J., & Koenderink, J. J. (2010). The shading cue in context. *i-Perception*, *1*(3), 159–178.

Chapter 23
Symmetry and Uprightness in Visually Perceived Forms

Lydia M. Maniatis

FROM PROJECTION TO PERCEPT

We live in a world of volumes and the spaces between them. Our perceived world shares this three-dimensional (3D) structure, though the splash of photons on our retina—the raw material for the visual process—contains no third dimension. The top of my coffee table appears to have near and a far side, a horizontal surface and four right angles; its retinal projection has no near or far side, no horizontal surface, no right angles. My mental representation of the coffee table is confirmed as I rest my feet upon it and reflect on how I knew. . . .

Veridicality (the accurate representation of the physical world in the mental) is a routine achievement of our visual system. Nevertheless, there exists a large class of objects with respect to which failures of veridicality are common. They are called pictures. Like the real coffee table, the flat picture of a coffee table may have an apparent near and far side, an apparently horizontal surface, and four apparently right angles. This organized misperception of the geometry of the pictorial surface may arise even if the figure consists simply of lines; such figures allow us to isolate, manipulate, and analyze the geometric determinants of the transformation from two-dimensional (2D) retinal projection to 3D percept.

Whenever a retinal projection (whether this derives from a 3D object or from a pictorial surface) is interpreted by the visual system as the projection of a surface that is not frontoparallel (i.e., not parallel to the retinal surface), then the retinal image will differ in shape from the source of the projection in (a) the sizes of its internal angles and/or (b) the relative extents of its surfaces. The latter differences arise because, when an extent is assumed to be receding, then it must also be assumed to have undergone foreshortening in the projection. In order to recover the true dimensions of the putative object, the visual system must undo this foreshortening by stretching the relevant dimensions in the percept. The former differences arise as the visual system must undo the projective distortions of the angles produced by receding surfaces.

If we are dealing with a typical object in the world, these shape adjustments will cause the mental representation of the object to resemble the actual object more than it resembles its projection. In the case of a pictorial 3D percept, however, these actions will result in a discrepancy between the mental representation of the object and the object itself, which is the 2D picture. By critically observing how apparent dimensions and angle sizes change as we vary their geometric context, we may infer some of the principles underlying and driving the projection-percept transformation.

RECEDING OBLIQUES AND THE SHEPARD ILLUSION

A classic demonstration of equivalent extents appearing unequal is Shepard's box illusion (Shepard, 1981; see Figures II.23-1a and II.23-1b). Figures II.23-1a and II.23-1b have identical tops. Each top is a rotated version of the other, such that the obliques in Figure II.23-1a are horizontal in Figure II.23-1b and vice versa. The two figures are not identical as wholes—in each case, a rectangle is attached to the horizontals and a parallelogram to the obliques. In other words, the rectangle is attached to a different pair of edges in each case.

The two differently oriented box tops appear to have different proportions, and this seems to stem from the fact that pairs of edges appear longer when they are oblique than when they are horizontal. These effects are consistent with our interpreting the 2D figures as projections of 3D box shapes. If the figures really were the projections of boxes,[1] then the obliques, which appear to recede away from the rectangular front face, would have to be interpreted as foreshortened and would therefore need to be perceptually stretched in order to recover the boxes' true proportions. Since a different pair of edges is stretched in each case, the two boxes differ in their perceived shape. This is essentially the explanation that Shepard (1981) gives for the illusion.

There is, however, another, confounding possibility. Deregowski and Parker (1992) have shown that we generally tend to perceive obliques in a picture as receding, even if they are attached to just one or two other lines and do not form part of an apparent surface or volume. If they are assumed to be receding, then they must be assumed to be foreshortened and must be perceptually stretched. Perhaps this orientation bias, not 3D geometric constraints, is the basis of Shepard's illusion.

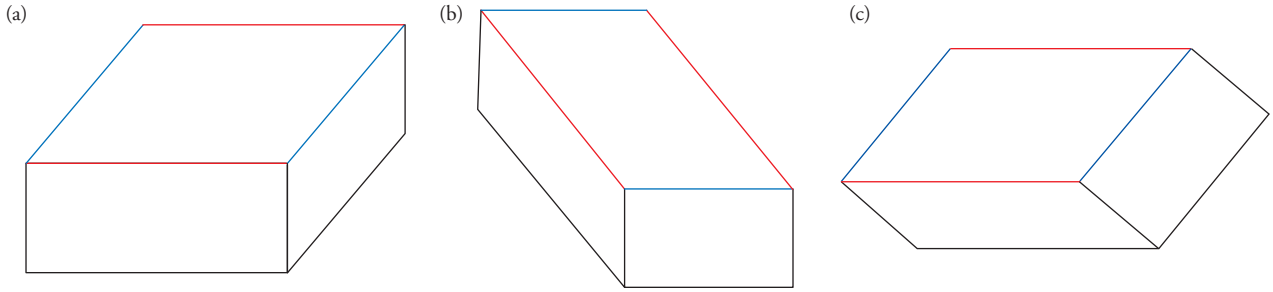


Figure II.23-1. All similarly colored lines are the same length. (From Maniatis, 2010a).

This assumption may be tested by adding a third box to the set (Maniatis, 2010a). Figure II.23-1c is simply Figure II.23-1b rotated, so that the obliques are horizontal and the horizontals are oblique. Unlike Figures II.23-1a and II.23-1b, which only have matching tops, Figures II.23- II.23-1b, and II.23-1c are identical as wholes. Nevertheless, the rotation of Figure II.23-1b to produce Figure II.23-1c instigates a change in the box's apparent proportions. Edges that were horizontal in Figure II.23-1b have become longer in Figure II.23-1c, where they are oblique; and the edges that were oblique in Figure II.23-1b have become shorter in Figure II.23-1c, where they are horizontal. With respect to projective geometry, there is no reason why Figure II.23-1c, being identical to Figure II.23-1b, could not be similarly interpreted as the projection of a box whose rectangular face, though tilted, is approximately parallel to the viewer's visual field and whose horizontal edges are receding and foreshortened. (In this case, the horizontals would also appear tilted downward). Interpreted in this way, the two boxes should look identical in shape. The fact that they do not seems to argue that, indeed, the perceived transformations are simply due to the orientation of the edges, à la Deregowski and Parker (1992), rather than being dictated by the geometry of the figure.

However, in this case the top of Figure II.23-1c should appear identical in proportion to the top of Figure II.23-1a, because the corresponding pairs of edges have the same orientation. In Figure II.23-1a, however, the blue edges appear approximately equal to the red edges, while in Figure II.23-1c the blue edges appear shorter than the red edges. These differences in proportion are consistent with a 3D shape–mediated determination of which pair of edges within the figure should be treated as more receding (and thus should be perceptually stretched) and which pair should be seen as frontoparallel (and thus not stretched), rather than with a determination based on the edge's relative orientation on the page.

Thus both orientation and shape constraints seem to be in play. If a figure's 2D shape were the only factor mediating the apparent length of an edge, then the proportions of the top of Figure II.23-1b should appear the same (though differently oriented) shape as the top of Figure II.23-1c. If the orientation of an edge were the only criterion, then the apparent proportions of the top of Figure II.23-1c should match the top of Figure II.23-1a. Instead, the final percept appears to be the result of an interaction between the geometry of a figure and its orientation on the page. Specifically, if both geometry and orientation call for an edge to be stretched, it will appear longer than if only one of the two criteria demand it, and if one of the two criteria demand it, it will appear longer than if neither criterion demands it.

We may observe this dynamic enacted in the blue lines in Figure II.23-1. In Figure II.23-1c, orientation—but not shape—calls for the blue edges to be stretched. They thus appear longer than in Figure II.23-1b, where neither criterion calls for stretching, but shorter than in Figure II.23-1a, where both shape and orientation considerations call for stretching. In the following section I attempt to show that the tendency of obliques to stretch is the result of a bias of the visual system to assume that objects are "upright."

RECEDING OBLIQUES, MIRROR SYMMETRY, AND UPRIGHTNESS

In Figure II.23-1c, the rectangular face is seen as more receding and less frontoparallel than is the case with the rectangles in the other two figures.[2] Seeing the rectangle as both right-angled and receding is not consistent with projective geometry, as a rectangle that is receding will not produce a right-angled projection. If the projection of a receding surface is rectangular, then the surface itself should not be perceived as rectangular. The lack of perceptual distortion of the 2D rectangle despite its apparent slant reflects an apparent resistance of rectangular projections to perceptual deformation. Non-right-angled quadrilaterals, such as trapezoids, on the other hand, do tend to exhibit the expected shape/slant interactions, so that a change in perceived slant corresponds to a change in perceived shape. This useful feature of trapezoids will help us to elucidate further the visual criteria that lead to perceptual stretching of obliques.

Figure II.23-2a is composed of a trapezoid, a rectangle, and a parallelogram. The trapezoid is symmetrical and appears (more or less) frontoparallel, so that the percept shares the shape and symmetry of the 2D form. The trapezoid also appears upright, since its axis of symmetry is aligned with the vertical of the page and of the visual field. The rectangle is oblique, but its oblique long edges (and those of the parallelogram-shaped roof) are interpreted as horizontal edges that are receding in depth, their obliqueness merely a trick of optics. With a vertical axis of symmetry and a base whose edges appear to bound a horizontal surface, the 3D object as a whole is perceived as upright. (For a more detailed discussion of perceived uprightness, see Maniatis, 2010b.)

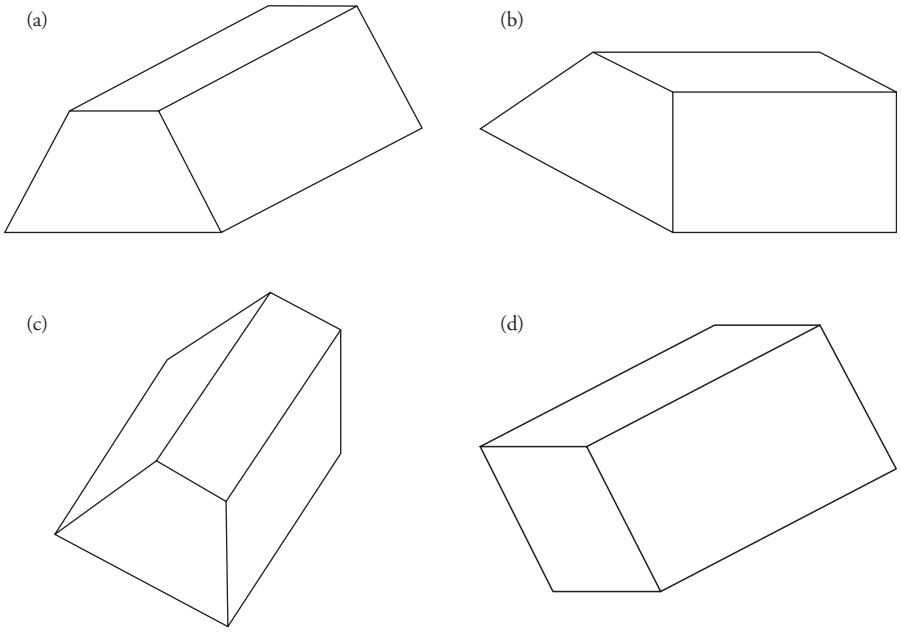


Figure II.23-2. The trapezoids in Figures II.23-2a, II.23-2b, and II.23-2c are identical in size and shape. The rectangles as well as the long, narrow parallelograms in Figures II.23-2a, II.23-2b, and II.23-2d are identical in size and shape. The lengths of the borders in the rectangle in Figure II.23-2c are identical to the length of the edges of the upper parallelograms in Figures II.23-2a, II.23-2b, and II.23-2d. (Figures II.23-2a, II.23-2b and II.23-2d from Maniatis, 2010a.)

In order to create a tilted version of Figure II.23-2a—that is, I want to manipulate the figure so that the trapezoid's axis of symmetry will appear tilted rather than vertical—I rotate the figure. To produce Figure II.23-2b, I rotated Figure II.23-2a just enough to align the long sides of the rectangle with the horizontal axis of the page. My manipulation does not give me the tilted-trapezoid percept I expected. The trapezoid no longer appears mirror symmetrical and frontoparallel but asymmetrical and receding. The changes in its shape and slant are mutually consistent, linked together by the principles of optical geometry. In other words, an asymmetrical shape such as the one we perceive in Figure II.23-2b, with the slant that we perceive it to have, would (more or less) produce a trapezoid-shaped projection on the page.

Why did a simple rotation of the figure within the plane produce these changes in the form's orientation in depth, and, consequently, its apparent 3D-shape interpretation? What was gained?

By reinterpreting the trapezoid as rotated in depth rather than as tilted, we have sacrificed its symmetry, but we have also "explained away" the obliqueness of its base (which forms the left side of the base of the whole figure) as the projective rise of an actually horizontal edge (as we did with the oblique long edges of the receding rectangle and parallelogram of Fig. II.23- II.23-2a). At the same time, the rectangular side now has a vertical axis of symmetry, and its base is aligned with the horizontal of the page. If it is perceived as frontoparallel (and it is, more or less), then its 2D mirror symmetry confers symmetry on the perceived volume as a whole. With a vertical axis of symmetry and a horizontal base, the perceived object is, once again, both symmetrical and upright. Thus, in Figures II.23-2a and II.23-2b, perceiving the obliques as receding results in symmetrical and upright percepts.

Note that because Figure II.23-2b is just a rotated version of Figure II.23-2a, it is just as consistent with a percept that looks like a tilted Figure II.23-2a—that is, a form with a trapezoidal cross-section, a tilted axis of symmetry, and a nonhorizontal (downward-tilted) base—as it is with the object we actually perceive. But our visual system does not let us see it this way, switching instead to the alternative that is both symmetrical and upright.

But we cannot always achieve both symmetry and uprightness. In Figure II.23-2c, the same trapezoid in the same orientation as in Figure II.23-2b (here I am referring only to the trapezoidal face, not to the whole figure) now appears symmetrical (and slanted in a direction consistent with symmetry both in the projection and in the percept) rather than asymmetrical. Why the switch? There is no way Figure II.23-2c can be reconciled with an underlying object that is both symmetrical and upright. The option of inferring symmetry alone is, however, available, though it must be combined with an assumption of object tilt. The option of uprightness in the sense of an asymmetrical object whose base is aligned with the horizontal plane is available in principle, but it is not preferred and impossible to picture.

Let us also ask what would happen if we changed the trapezoidal front face of Figure II.23-2a into an asymmetrical parallelogram while keeping the rest of the figure intact. Would this new face continue to appear frontoparallel? If it did, it would mean that the whole figure would be seen as having an asymmetrical face and would be asymmetrical as a whole. (The corner view of the rectangle cannot, in principle, be reconciled with symmetry in 3D.) We explore this question with Figure II.23-2d.

The altered face does not remain frontoparallel but appears to recede along its horizontals. Because they appear receding, these horizontals are assumed to be foreshortened in the projection and are stretched. This affects the shape of the parallelogram that forms the roof of the figure (and shares one of its horizontals with the front parallelogram); it appears wider than its (identical in shape and orientation) twin in Figure II.23-2a. Meanwhile, the rectangular side now appears much more nearly frontoparallel than its Figure II.23-2a counterpart.

Why does the rectangle appear more nearly frontoparallel in Figure II.23-2d than in Figure II.23-2a? Now that the trapezoid is a parallelogram, the only way we can

presume symmetry in the object is to treat the rectangle as frontoparallel, so that its 2D symmetry is consistent with symmetry in the perceived object. Consequently, the obliqueness of its long edges—one of which it shares with the roof—may no longer be "explained away" as projective rise of receding, and foreshortened, horizontals. These edges are thus not (or are less) stretched, and the object is perceived as tilted. As a result of the various changes in foreshortening assumptions, the roof of the figure appears not only wider but also shorter than its (identical in shape and orientation) Figure II.23-2a counterpart—rather more like its counterpart in Figure II.23-2b.

To summarize the lessons of Figure II.23-2: Figures II.23-2a and II.23-2b, identical except for their orientation on the page, offer two symmetrical 3D solutions—one with a trapezoidal and one with a rectangular cross-section. Each choice embodies different depth relationships among the parts of the figures and requires corresponding shape changes. Despite the availability of the two solutions, the two figures are not bistable. In each case, the shape percept that is realized is the one that results in a more upright object. In Figures II.23-2c and II.23-2d, there is no solution that is both symmetrical and upright, so that the tilted trapezoidal or rectangular side, respectively, becomes more frontoparallel and the percepts are symmetrical but tilted. An attempt to maximize uprightness is also consistent with the Shepard box of Figure II.23-1c (which is similar to Fig. II.23-2d), where the rectangular side does not appear as frontoparallel as its twin in Figure II.23-1b. If it did, then the rectangle (and consequently the whole perceived volume) would have to appear much more tilted (because none of the tilt could be explained as projective rise of receding edges).

In general, it appears that obliques recede when, by doing so, they will contribute to a more upright and more regular whole-figure. In Figure II.23-2b, the top and bottom edges of the trapezoid recede, spoiling the symmetry of this face but not of the whole, which becomes upright and symmetrical around the rectangle. The same edges in the identically oriented trapezoid of Figure II.23-2c do not recede, since, here, treating the trapezoid as symmetrical is the only road to a symmetrical object. The oblique sides of the trapezoid in Figure II.23-2a, though oblique, do not appear to recede, as neither bilateral symmetry nor uprightness would be thereby enhanced.

WHY SHOULD THERE BE AN ORIENTATION CONSTRAINT?

In the natural world, the majority of objects—most notably living things—are mirror symmetrical, and their axis of symmetry tends to be aligned with the direction of gravity. For these reasons, the vertical of our retinas will typically be aligned with the direction of gravity as well as with the viewed object's axis of symmetry. In addition, we actively take steps to find a viewpoint on an object such that its projection has an axis of symmetry (if one is available) and to align this axis with our retinal vertical; we see this in the way we will rotate an object in our hand in order to inspect it or irresistibly straighten a picture on a wall.

If mirror symmetry and uprightness are typical properties of objects viewed from a typical viewpoint, then it makes sense for the visual process to try to reconcile each projection with these properties. Prior assumptions such as these may be supplemented by additional processes, such as those that compare the simultaneous images in the two eyes to detect binocular parallax or successive images in one eye to detect motion parallax. This information may be compared to the depth relationships assigned on the basis of prior assumptions and potential discrepancies resolved. The power of the shape and orientation assumptions, however, is proven by our irresistible tendency to see 3D objects in pictures when these are "better" than their 2D projections, even though the complete absence of parallax within them is an unambiguous sign of flatness.

NOTES

1. In fact, the Shepard boxes are not perfectly consistent with projections of boxes. As Pizlo (2008) has observed: "The visual system can tolerate some inconsistencies between a percept and the available 2D retinal image. It tolerates this inconsistency in order to preserve the symmetry and compactness of the 3D shape percept" (p. 156). This should be kept in mind in the discussion that follows.
2. While for clarity I have restricted my discussion of structural biases to symmetry, this factor may in most cases be correlated with others, in particular with compactness.

REFERENCES

Deregowski, J. B., & Parker, D. M. (1992). Three-space inference from two-space stimulation. *Perception & Psychophysics*, *51*(4), 397–403.

Maniatis, L. M. (2010a). Alignment with the horizontal plane: Evidence for an orientation constraint in the perception of shape. *Perception*, *39*(9), 1175–1184.

Maniatis, L. M. (2010b). Shape, gravity and the perception of the right angle. *Perception*, *39*(9), 636–640.

Pizlo, Z. (2008). *3D shape*. Cambridge, MA: MIT Press.

Shepard, R. N. (1981). Psychophysical complementarity. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 279–341). Hillsdale, NJ: Lawrence Erlbaum.

Chapter 24
The Bathtub Illusion

Lydia M. Maniatis

GUESSING AT THE WORLD BEHIND THE RETINAL PROJECTION

Pictures are puzzles. Though obviously flat, they can recreate, in our minds, all of the structural complexity of the three-dimensional (3D) world. In this respect, however, they are no less—or more—puzzling than the projections objects in the world cast on the surface of our retinas. These projections are also flat, and they also create 3D mental images. In other words, whether the light is coming from a picture or from a 3D scene, the problem for the visual system is the same—what kind of objects in the world caused the two-dimensional (2D) retinal projection?

The answer to this question is, in every case, a guess. Every visual image is a guess, as each retinal projection could have been caused by an infinite number of real-world structures. The guess is informed by evolutionary adjustments of the visual process that have tended to produce a good match between scene and percept in our particular environment. These adjustments may be described as assumptions that the visual system makes about the relationship between the retinal projection and the real world. If, for example, a projection is consistent with both geometrically regular and geometrically irregular 3D objects, our guesses will tend to favor the more regular possibilities (see, e.g., Li, Pizlo & Steinman, 2009). Thus a retinal projection that is consistent with a cube tends to produce a cube percept, even if this projection comes from a flat surface in the world (a picture), because the flat projected shape is usually less regular than the 3D cube structure. If the 2D projection happens to be as regular as the 3D option, as if a cube casts a hexagonal projection, then we will tend to perceive the 2D version (Kopfermann, 1930, cited in Koffka, 1935).

As we move around the cube, its projection changes shape, but the real-world cube does not appear to change because each new projection is best resolved by the same, cube-shaped "guess." All that needs to change to reconcile the projections with the cube shape is our apparent viewpoint. This is typical of most of the objects that we see in the world; our apparent mutual orientations change while the object's perceived shape stays the same.

PROJECTIONS OF PICTURES OF 3D OBJECTS VERSUS PROJECTIONS OF 3D OBJECTS

This is, however, often not the case with pictures of 3D objects. The shapes of the projections of 3D objects tend to be weird—as is evident if, for example, we trace the outlines of a picture of a cube. As we move around the picture, the new projections produced by these weird shapes are not the same as those the real cube would produce, and they will not necessarily be compatible with cube-shaped guesses. The guesses they produce will depend on the shape of the projection and the best (e.g., most regular) guess that the visual system can come up with for that particular projection.

The bathtub photo shown in Figure II.24-1 is a particularly striking example of how changes in the retinal projection produced by changing our viewpoint can cause changes in the shapes of the perceived objects. Figure II.24-1a shows the picture viewed head-on; Figure II.24-1b is the view from the left; and Figure II.24-1c the view from the right. Each new viewpoint (i.e., each new projection) produces a percept with a different shape and orientation. All of the projections are consistent with a fairly regular bathtub, as we can see, but not with tubs having the same proportions as the original. The change in proportions is largely due to changing assumptions about which lines have been foreshortened in the projection/picture. This, in turn, is linked to changes in the mirror symmetry of the projected visible faces of the bathtub and to the degrees of obliqueness of their edges (for a discussion see chapter 23 in this volume).

In the left view, shown in Figure II.24-1b, the projection of the left face of the tub is more symmetrical than the right, and its upper and lower edges are more aligned with the horizontal of the picture plane. Both of these factors contribute to the left face being perceived as more frontoparallel than the right face, whose 2D asymmetry and oblique upper and lower edges are both consistent with a surface that is symmetrical in reality and receding along these edges. Edges that are perceived as receding more are assumed to be more foreshortened in the projection and are consequently stretched more in the percept so as to recover the dimensions of the putative object. Edges perceived as more frontoparallel/less receding are less stretched. The



Figure II.24-1. Photo of a bathtub seen from different vantage points. (a) Face-on. (b) View from the left side. (c) View from the right side. Photo by Lydia M. Maniatis.

upper and lower edges of the right face in Figure II.24-1b are, consequently, stretched more than those of the left face. Exactly the reverse occurs in the case of right view shown in Figure II.24-1c, where the left face becomes the more asymmetrical, with the more oblique upper and lower edges. As a result, the two views to produce tubs with very different proportions. The relationship between Figures II.24-1b and II.24-1c is similar to that between the Shepard boxes discussed in chapter 23 in this volume (see Figures II.23-1a and II.23-1b). Here too edges are perceived as longer or shorter when more or less oblique and bound to more or less symmetrical surfaces.

That pictured objects may change shape and/or orientation with changes in vantage point is not a new observation but rather one that has perennially engaged theorists (recent analyses include Papathomas, Kourtzi, & Welchman, 2010; Todorovic, 2008). Erkelens (2013), using simple grid stimuli that made it easy to estimate the shape of the retinal projection with changing vantage point, concluded that what is perceived when we view a picture from various angles is a direct function of (and thus could be predicted based on) the shape of its retinal projection. That is, if a retinal projection from a frontoparallel surface produces percept *x*, then a similar retinal projection from a slanted surface will produce a similar percept, even under binocular viewing.

In contrast, Vishwanath, Girshick, and Banks (2005) found that under binocular, but not monocular, viewing observers corrected for the slant of the screen, so that even when viewing the picture obliquely, they saw shapes corresponding to the frontoparallel projection. The difference between these two studies may stem from the type of pictorial stimuli used. As observed, for example, by Sedgwick (1993), our perception of pictures appears to be a compromise between aspects of the drawing consistent with it being the projection of a 3D scene and information (e.g., binocular and motion parallax) indicating that the drawing is a feature of a flat surface.

The compromise is detectable in Erkelens's (2013) data, in which the perceived slant of a grid was less than it would be if the convergence of edges in each 2D retinal image were completely "undone" to produce a perceptually regular grid. Some of the stimuli used by Vishwanath et al. (2005) are arguably weaker than those of Erkelens in pushing the 3D "argument." They consist of projections of a sphere, a 3D ellipse, and a cube with only two visible sides. Even in frontoparallel viewing, a particular shape, for example a parallelogram, may appear more or less slanted (and, in the case of the parallelogram, more or less right-angled) depending on the whole shape in which it is embedded. For example, the famous Shepard tabletop illusion (Shepard, 1981) is more convincing when the planes are embedded in box shapes than when they are presented in isolation. In other words, even in the context of frontoparallel viewing, the cues to the flatness of the picture plane can be more or less dispositive. Vishwanath et al. drew a

similar conclusion in the case of a grid stimulus similar to the ones used by Erkelens. They had observers report on the perceived aspect ratio rather than the slant of the stimulus, but it is not clear why this should produce the observed differences between the two studies.

WHY PICTURES DON'T OFTEN SEEM TO CHANGE WITH CHANGING VIEWPOINT

What is clear is that pictures can and do change shape with changing viewpoint. Why are such changes so rarely observed on casual viewing? First, as discussed, cues to flatness of the surface may be more or less influential. Second, and more importantly, if a shape is regular on the surface of a page—if, for example, it is a rectangle—then, just as in the case of most other regularly-shaped objects, each new retinal projection will be best resolved as that same shape, with apparent (and actual) slant relative to the viewer changing in tandem with the apparent (and actual) slant of the picture surface. Third, if we are looking at a pictured object from relatively far away, the changes in our viewing angle and in the retinal projection as we move around will be fairly small. Finally, even fairly large changes in the shape of the percept may not be noticed as they occur in real time. This was noticed by a viewer of a video version of the bathtub effect. The picture in the video is slightly different from that of Figure II.24-1, and the effect not quite as dramatic, but it is still substantial. Several of the YouTube commenters for this video complain that they don't get it (and the illusion has received more dislikes than likes). One commenter, however, makes the relevant observation:

> If you stop play, then jump to the start and end of the clip back and forth I actually see what you mean! But I think I can't figure it by seeing the clip in normal play. I mean I don't see it as something unusual. But if I look at the start and end pictures there is surely something. Too bad one cannot scrub through the youtube timeline. But one can scrub oneself in the bathtub though.

REFERENCES

Erkelens, C. J. (2013). Virtual slant explains perceived slant, distortion, and motion in pictorial scenes. *Perception*, *42*(3), 253–270.

Koffka, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt, Brace.

Kopfermann, H. (1930). Psychologische Untersuchungen uber die Wirkung Zweidimensionaler darstellungen korperlicher Gebilde. *Psychologische Forschung*, *13*, 293–294.

Li, Y., Pizlo, Z., & Steinman, R. M. (2009). A computational model that recovers the 3D shape of an object from a single 2D retinal representation. *Vision Research*, *49*(9), 979–991.

Papathomas, T. V., Kourtzi, Z., & Welchman, A. E. (2010). Perspective-based illusory movement in a flat billboard—An explanation. *Perception*, *39*(8), 1086–1093.

Sedgwick, H. A. (1993). The effects of viewpoint on the virtual space of pictures. In S. R. Ellis (Ed.), *Pictorial communication in virtual and real environments* (2nd ed., pp. 460–479). London: Taylor & Francis.

Shepard, R. N. (1981). Psychophysical complementarity. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 279–341). Hillsdale, NJ: Lawrence Erlbaum.

Todorovic, D. (2008). Is pictorial perception robust? The effect of the observer vantage point on the perceived depth structure of linear perspective images. *Perception*, *37*(1), 106–125.

Vishwanath, D., Girshick, A. R., & Banks, M. S. (2005). Why pictures look right when viewed from the wrong place. *Nature Neuroscience*, *8*(10), 1401–1410.

Chapter 25
The Pitchroom Illusion

How High Is Up?

Leonard Matin, Ethel Matin, Wenxun Li, Todd E. Hudson, and Adam Shavit

A BIT OF HISTORY: CURARE AND THE OCULOPARALYTIC ILLUSION

In 1974 John Stevens, who was then a graduate student at the University of Pennsylvania, told one of us (LM) about his as-yet-unpublished experiments (Stevens et al., 1976), in which he had undergone transient experimental paralysis in the interest of determining whether inflow or outflow mechanisms were involved in maintaining perceptual stability when voluntary eye movements are made or attempted. When John told me about his work, I asked him whether it would be possible for me to undergo paralysis in his setup, and he responded with an invitation to visit. A postdoctoral scientist (Evan Picoult), a graduate student (Rodger MacArthur), and an undergraduate (David Young) and I traveled to Penn and met with John and the anesthesiologist (McIver Edwards). We brought with us some primitive but flexible equipment that we had constructed at Columbia for generating visual and auditory stimuli for measuring spatial localization.

We discovered that weakening the extraocular muscles under partial paralysis reduced the ability of the extraretinal eye position signal to compensate for changes of the position of the eye in dose-dependent fashion. The error in perceived elevation of a luminous target increased with elevation of the eye in the head from a no-illusion orientation (Matin et al., 1982; Matin, Stevens, & Picoult, 1983). *But, most important, this mislocalization only occurred in darkness.* When room illumination was restored to the normally oriented, structured visual environment of the laboratory, the mislocalization did not occur. The prediction that visual pitch would generate similar errors in observers for whom extraretinal compensation was normal was a short intellectual leap. Fueled by the desire to produce the equivalent of this oculoparalytic illusion that anyone could experience without paralysis, we were led to the construction of the pitchroom and to the discovery of the pitchroom illusion.

INFLUENCE OF VISUAL PITCH ON PERCEIVED ELEVATION

Basics of the Pitchroom Illusion

Egocentric spatial illusions are perceptual mislocalizations of objects in space or of the observer's body or its parts. They have several significant properties: (a) with appropriate settings of the illusion-inducing stimulus, the effects are large; (b) they involve the visual system and also the vestibular, somesthetic, and proprioceptive systems; (c) they can be manifested as mislocalizations of objects, of whole environments, and/or mislocalizations of the felt position of the body or of its parts (e.g., the head, eye, hand, arm, leg); and (d) they generally manifest not only as perceptual mislocalizations but as peculiarities of manual behavior and other spatial peculiarities related to the body.

The egocentric illusion that we describe in this chapter involves perceptual mislocalizations of elevation (hence the subtitle, "How High Is Up?"). We refer to it as the pitchroom illusion, a reference to the instrument (Fig. II.25-1a) employed in the initial studies of the illusion (Matin & Fox, 1986, 1989, 1990). However, very similar although somewhat attenuated effects can be obtained with two long eccentrically located lines (Fig. II.25-1b; Matin & Li, 1992a, 1992b) or even with one line in either the left (Fig. II.25-1c) or the right (Fig. II.25-1d) visual field (Matin & Li, 1992a, 1994c). In all cases, when the inducing stimulus is pitched topbackward (as in Figures II.25-1a, II.25-1b, II.25-1c, and II.25-1d), a visual target at true eye level appears higher than true eye level; when the field is pitched topforward, a visual target at true eye level appears lower than true eye level. The inducer's influence is very large. For example, erect observers in the pitchroom locate the elevation of their own eyes anywhere from below their waists up to several feet over their heads as measured by the physical setting of the elevation of a small luminous target projected on the inside rear wall of the pitchroom. The magnitude of the effect depends systematically on the magnitude and

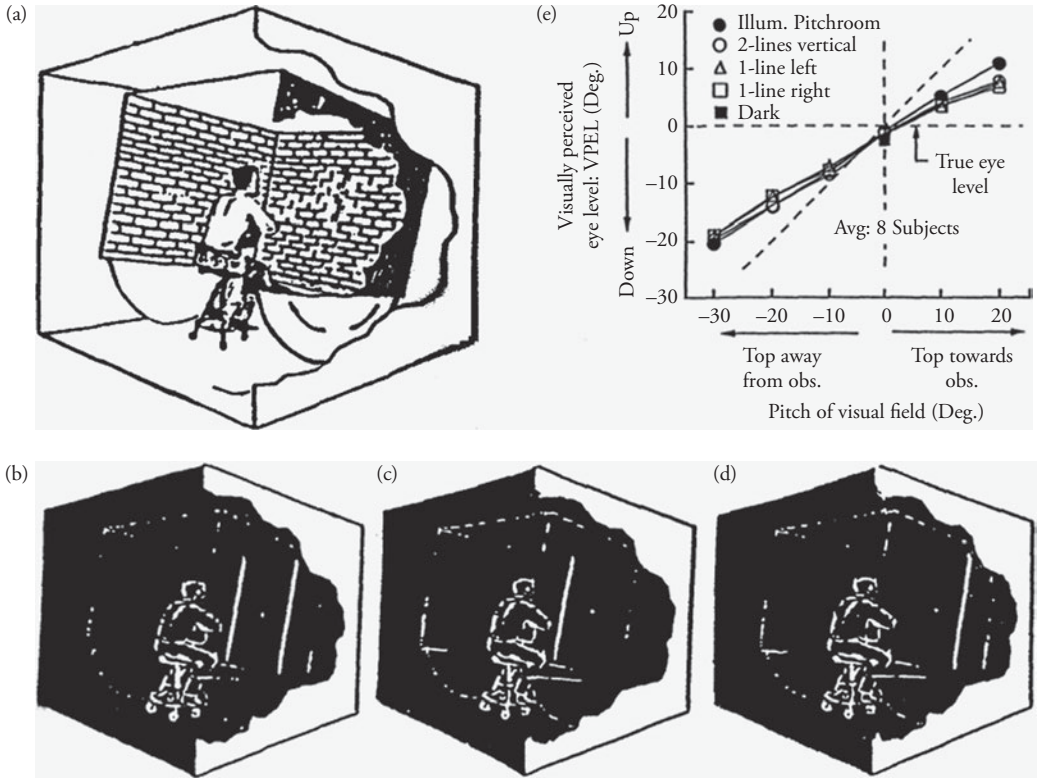


Figure II.25-1. Visual inducer and pitchroom illusion. (a) The pitchroom in its normally illuminated condition with the complexly structured visual field viewed by the seated observer. (b) The pitched-from-vertical 2-line inducer. (c) The left pitched-from-vertical 1-line inducer. (d) The right pitched-from-vertical 1-line inducer. The visual target is represented as a small white point in the center of the field in (a)–(d). (e) Pitchroom illusion: VPEL systematically changes with the pitch of the visual inducer in (a)–(d). In (e), the dark value is displayed below the abscissa zero. (Data from Matin & Li, 1994c.)

direction of the room's visual pitch and on the distance of the observer from the wall.

In addition to the effects of inducer pitch on perceived elevation, several other effects are associated with viewing the pitchroom. First, there are large influences on the perceived size and perceived pitch of objects (Hudson, Li, & Matin, 2000; Matin & Fox, 1986, 1989). For example, when an erect observer views a person who is standing in front of the topbackward rear wall, the person appears to be considerably taller than life-size and to be pitched forward; the observer generally also observes some distortions of the body, with the top appearing to have increased in size relative to the bottom. With topforward pitch, the viewed person appears to be standing in a hole and/or to be pitched backward. Second, a plumb line hung from the pitchroom's ceiling is perceived as pitched in the direction opposite to the pitch of the room. Third, when the pitchroom is rocked back and forth, the elevation of a stationary, luminous target projected against the rear wall at a fixed height appears to change synchronously in the direction opposite to the motion of the pitchroom. Fourth, a ball or water in a tube is perceived to be rolling uphill. Finally, some observers report feeling dizzy.

Psychophysical measurements and observations have been made on related illusions in natural environments (MacDougall, 1903; O'Shea & Ross, 2007; Stoper, 1998). Most of the phenomena experienced in the pitchroom also occur in some natural settings. The Mystery Spot at Santa Cruz, California, where visitors report plumb lines that hang askew as if held by unseen forces, and errors in the perceived size of people is well known. Because the visitors do not know that visual features of the environment are creating illusions, the owners of the site are able to market it as a "gravitational warp" or a "magnetic disturbance" that has baffled scientists.

The Observer's Task, VPEL, and the Linear Weighted Averaging Model

The inducing effects have been measured by using observers' reports about a visual target's height relative to their eye level (too high, too low) in order to allow the experimenter to obtain a null point for visually perceived eye level (VPEL). The VPEL error is one point in an essentially undistorted translation of the entire scale of elevation (Robison, Li, & Matin, 1995).The basic data in studies of the pitchroom illusion are the functional relationship between VPEL and inducer pitch. As can be seen in Figure II.25-1e, the inducer's effects are very large and vary linearly with pitch over a wide range. The slope of the function provides an overall measure of the inducer's effect: A slope of 1.00 (dashed diagonal line in Fig. II.25-1e) would imply that the visual inducer was the only factor determining the observer's perception of elevation. A slope of 0.00 would imply that the inducer played no role. The literature shows that the inducer plays a large but not exclusive role (Matin & Fox, 1989; Matin & Li, 1994c; Poquin,

Ohlmann, & Barraud, 1998; Post & Welch, 1996; Stoper & Cohen, 1989; Welch & Post, 1996). A number of these other sources of influence (e.g., gravity) have been explored in some depth (Chelette, Li, Esken, & Matin, 1995; DiZio, Li, Lackner, & Matin, 1997; Li, Dallal, & Matin, 2001; Matin & Li, 1995a, 1995b). We have lumped these influences together and labeled their combination the "body-referenced mechanism" (Matin & Fox, 1989). A treatment employing a linear weighted average of influences from the visual field and the body-referenced mechanism was developed to represent the data (Matin & Fox, 1989; Matin & Li, 2001, 2010):

$$\text{VPEL} = k_v V + k_b B$$

Here V represents the influence of the visual field, B represents the influence of the stimulus basis for the body-referenced mechanism, and k_v and k_b are weighting constants for the visual field and the body-referenced mechanism, respectively, with $k_v + k_b = 1$. For the results in Figure II.25-1e with a slope of 0.63, k_v and k_b equal 0.63 and 0.37, respectively.

THREE COMMON SPATIAL CUES THAT DO NOT INFLUENCE VPEL

This section describes experimental aspects of the pitched inducer that have played some role in other spatial phenomena and a priori might have seemed critical for VPEL: binocular stimulation, visual field complexity, and the perception that the visual field (the inducer) is pitched. None of the three was critical. Instead, as is shown in the next section, the critical aspect of the inducer is its *retinal orientation*.

Binocular Viewing Is Not Critical

Pitching the inducer produces changes in the cues to distance (e.g., with topbackward pitch, the top of the inducer is farther from the observer than the bottom; with topforward pitch, the top is closer to the viewer than the bottom). This raises the possibility that such cues along with the main stereocue of horizontal binocular retinal disparity might play a significant role in the induction effect. However, no differences were found between the VPEL measurements with monocular and binocular viewing (Matin & Fox, 1989; Matin & Li, 1992b; Stoper & Cohen, 1989).

A Complex Visual Field Is Not Critical But the Amount of Stimulus Matters

The dim (phosphorescent, <0.1 ml), thin, parallel 2-line inducer (Fig. II.25-1b) or the 1-line inducer (Figures II.25-1c, II.25-1d) presented in otherwise complete darkness produce effects on the VPEL versus pitch functions (Fig. II.25-1e) that are very similar to the function with the pitchroom: The effects with the 2-line and 1-line inducers are 88% and 82% of the effect with the pitchroom, respectively (Matin & Li, 1994c). Experiments that involved still further reduction of the stimulus to short lines and to arrays of points show that the inducer's effect (slope of the VPEL vs. pitch function) increases exponentially with the "amount of stimulus" approaching an asymptote of approximately 0.6, the slope in the pitchroom. The experiments showing these latter effects are further described in the later sections on lines, points arrays, and array orientation.

The Perceived Visual Pitch of the Inducer Is Not Critical

The close relationship between VPEL and the physical pitch of the surface containing the inducer led several workers to suggest that VPEL is determined by the pitch of an implicit surface (a mental representation of the plane surface generated by the inducer). This was investigated directly in experiments in which subjects made VPEL judgments and also made manual matches to the perceived visual pitch of the surface in the presence of a long, roll-tilted, parallel 2-line configuration within a frontoparallel plane on one side of the visual field in otherwise total darkness. As might be expected, the subjects' perceived pitch of the rolled inducer did not deviate significantly from the erect orientation for any of the roll-tilts. However, their VPELs were systematically changed with roll-tilt, in direct contradiction to the implicit surface view (Li & Matin, 1998). Further experimental work with measures of the perceived visual pitch of the inducer along with measures of VPEL in the fully illuminated pitchroom uncovered the surprising fact that the two are statistically independent: when the confounding effect of physical pitch was removed, perceived visual pitch and VPEL were orthogonal (Hudson et al., 2000; Hudson, Li, & Matin, 2006). Additional evidence of the independence of VPEL from perceived visual pitch was obtained in a study with a person (DF) who suffered extensive damage to the ventral visual stream but had an intact dorsal stream (Servos, Matin, & Goodale, 1995). Her VPEL reports under the pitched inducer's influence were similar to those of a normal control subject. However, she was not able to match the pitch of the inducer by manually adjusting the pitch of a plaque mounted on a pivot stand: her inability to "see" the pitch did not affect her ability to make VPEL judgments. Post, Teague, Welch, and Hudson (2003) also report that VPEL and perceived pitch are processed independently and suggest some contribution of physical pitch to VPEL beyond retinal orientation.

THE RETINAL ORIENTATION OF THE INDUCER *IS* CRITICAL

Work with inducers consisting of either one or two pitched or roll-tilted lines has shown that the *retinal orientation* of the inducer, not its physical orientation or its perceived orientation, is the critical factor in its effect on VPEL. In this section we provide the basic support for this conclusion and describe further work to clarify its importance and mode of operation.

The Evidence for Retinal Orientation and VPEL

The two individual lines in Fig. II.25-1b are parallel and vertical when erect (hence "pitched from vertical" when pitched). The two are bilaterally symmetrical relative

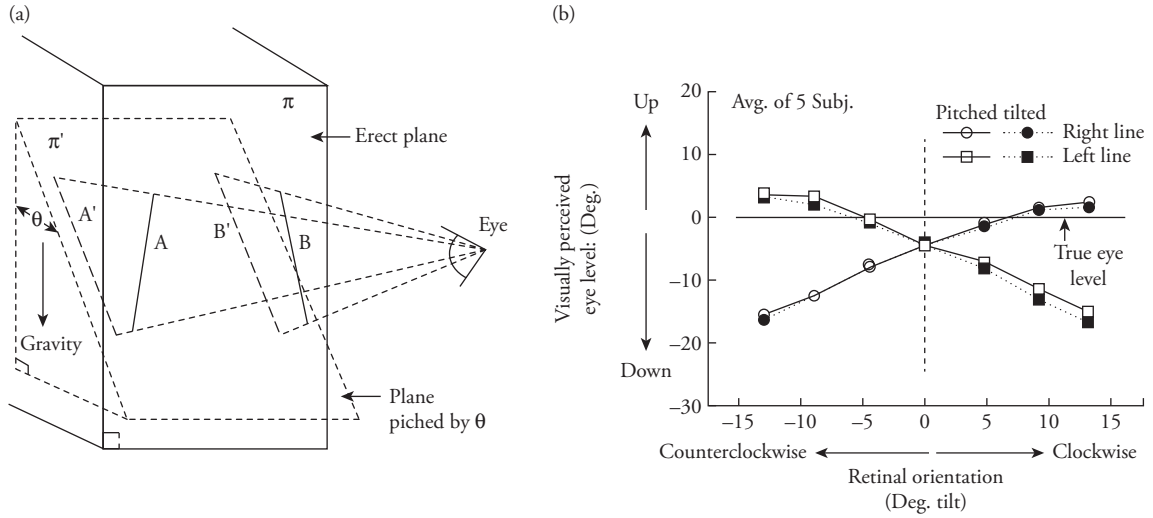


Figure II.25-2. Identical VPEL results are induced by lines with the same retinal orientation from pitched and erect planes. (a) Identical visual stimulation can be produced from the pitched-from-vertical line inducers (A' and B' on pitched plane π') and from the oblique line inducers (A and B on the frontoparallel plane π'). Both nodal planes pass through the nodal point of the eye. (b) VPEL is plotted against retinal orientation for each of the four 1-line inducers in (a). (Adapted from Matin & Li, 1994a.)

to the midsagittal plane of the observer. That it is their retinal orientation, and not the physical orientation in the distal stimulus, that is critical for their influence on VPEL is made clear by the results in Figure II.25-2b (Li & Matin, 1996; Matin & Li, 1994a) which show that a roll-tilted line in a frontoparallel plane exerts the same influence on VPEL as does a pitched-from-vertical line from a pitched-only plane that strikes the same retinal orientation. (Pitching a plane with horizontal lines produces only a small effect that is due to the height in the field, not the retinal orientation of the lines per se; Matin & Li, 1992a, 1992b; Post, Welch, & Clark, 2000.)

Line Length and Line Orientation

Figure II.25-3a shows the results of experiments in which length of the 1-line and 2-line inducers were varied. The slopes of the VPEL versus inducer pitch functions from these experiments, along with the pitchroom slope, are segments of a negatively accelerated exponential function of total length of all lines approaching an asymptote approximating 0.6 (Fig. II.25-3a). The function representing 1-line and 2-line results are indistinguishable, and the pitchroom result treated as an 8-line combination is very close to the asymptote.

Summation and Averaging Among Lines of the Same or Different Orientations

The significance of line length and line orientation is further elucidated in Figures II.25-3b and II.25-3c (Matin & Li, 1999, 2000). Each point plots the VPEL for a 2-line inducer (two lines located bilaterally symmetrically) on the ordinate against the sum of the two individually measured 1-line VPELs for the same two lines (abscissa). Each of the 49 pairs in a panel consists of one of the 49 2-line combinations of seven different 1-line orientations (–30° to +30°). Figure II.25-3c displays the results of an experiment for short lines (12° long) where the best fit slope of the line comparing the simultaneously viewed and sum of the separately viewed inducers is 0.99 where 1.00 would be the "complete algebraic summation" result. Figure II.25-3b displays the result for the same experiment with long lines (64° long) where the best fit slope is 0.61, significantly different from 0.5 (that would be predicted by "averaging two independent processes") but far from complete summation. These two different outcomes with different lines lengths indicate two of a range of possibilities for neural combination across the two hemiretinas.[1]

POINTS ARRAYS AND ARRAY ORIENTATION

Points Arrays

Any subset of points within a given straight line defines the same orientation as any other subset. Thus, in a further analysis of the inducer's influence on VPEL, several properties of the straight line have been examined in experiments with linear arrays of points. The number of equally spaced points, the spacing between points, and the length of points arrays have been varied systematically along with the orientation of one or two arrays of points. The slope of the VPEL versus pitch and VPEL versus roll-tilt functions increase exponentially with array length and with the number of points, reaching an asymptote with 25 points at an array length of 64° that equals the slope of the continuous line (Li & Matin, 1992; Matin, Li, Li, & Shavit, 2006; Matin, Matin, & Li, 2016).

Array Orientation

Two different sorts of orientation can be distinguished in a straight array of parallel lines: the orientation of the array ("array orientation") and the orientation common to the parallel individual lines in the array (the "line orientation").

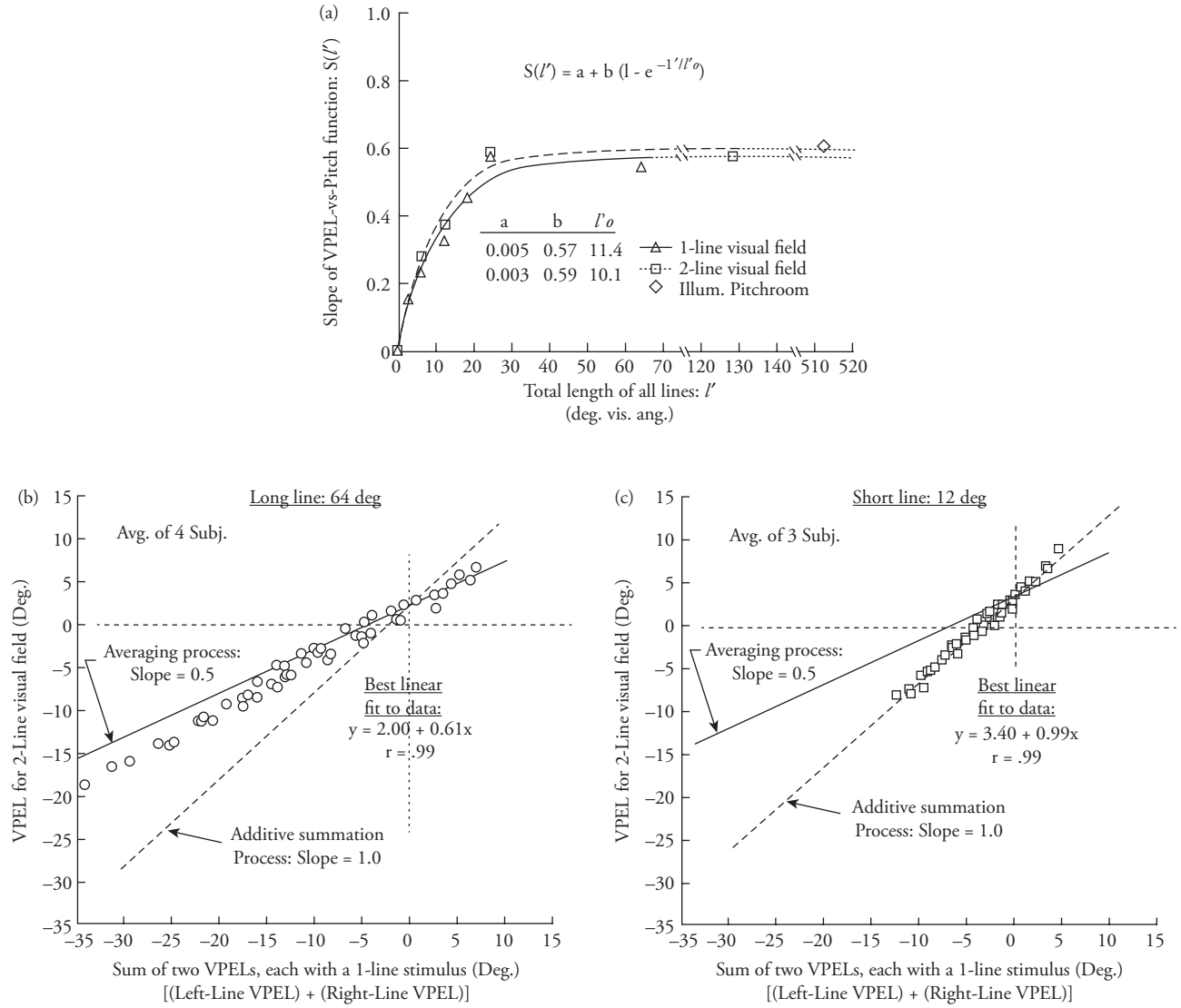


Figure II.25-3. Summation and averaging in the pitchroom illusion. (a) Using the total length of all lines as the common independent variable (abscissa), we display the pitchroom illusion as the slope of the VPEL versus pitch function (ordinate). The results for the 1-line and 2-line inducers and the illuminated pitchroom are well fit by the same negatively accelerated exponential. Panels (b) and (c) display 49 data points for short-line (12°-long) and long-line (64°-long) inducers, respectively. For each point the ordinate plots the average VPEL for a 2-line inducer against the sum of the two VPELs for the two constituent 1-line inducers measured separately. Each of the 49 different line pairs consisted of two orientations chosen from the same seven possible orientations ($7 \times 7 = 49$ pairs). The slopes of the two theoretical straight lines in each panel would result if the total visual influence contributing to the 2-line VPEL equaled the algebraic sum of the contributions of the two constituent lines (slope of 1.00, "additive summation") and the average of the VPELs of the two constituent lines ("averaging process"), respectively. The *y*-intercept and slope of the best-fitting straight line (not shown) in panels (b) and (c) are given by the equation. (Long-line results and short-line results are from Matin and Li, 1999 and 2000, respectively.)

The array orientation is the implicit line through the centers of the individual lines in the array and can be varied independently of the line orientation. In one set of experiments with short individual lines and long arrays, VPEL increased monotonically with array orientation (Shavit, Li, & Matin, 2004). In a second set of experiments in which array length and line length were varied along with number of lines, VPEL changed monotonically with line orientation and with array orientation. Line length and array length further modulated the influence of orientation (Shavit, Li, & Matin, 2008).

The Hand-to-Body Distance Effect

When observers in a pitched visual field point to the perceived elevation of a target with the unseen arm/hand, they point in directions that depend systematically on the distance of the hand from the body: they make gross errors when the hand is next to the body but point accurately at the perceptually mislocalized target when the arm is fully extended (Li & Matin, 2005a). Further studies showed that the distance effect disappears shortly after the inducer is extinguished, followed by a further slow exponential decay

(Li, Matin, & Matin, 2013). The slow decay of the manual effect was similar to the decay previously shown for VPEL (Matin & Li, 1995a).

The basic facts are illustrated in Figure II.25-4, which shows the effects of two inducer pitches (–30° and +20°) with various combinations of three distances (0 cm, 40 cm, and 70 cm) and two manual behaviors—(a) pointing and (b) setting the height of the arm to match the perceived elevation of an eye-level target (height matching). The dramatic distance effect (the large differences between the manual behavior at various hand-to-body distances while the inducer is on) collapses shortly after inducer offset. Additional work measuring the inducer's effect on the felt position of the arm (no visual target) showed similar short-lived induction effects. We conclude that the visual inducer has a short-lived effect on the felt position of the arm and on egocentric perceptual judgments, as has been shown for motor representations in the sensory-motor interaction literature (reviewed in Rossetti, 1998).

The distance effect is not limited to studies of perceived elevation; it was also found in studies of the classical rod-and-frame effect (Li, Matin, Bertz, & Matin, 2008). This generality suggest that it has some survival advantage. For example, the fact that that there is virtually no manual illusion when the hand is next to the body is important. If this were not the case, it might lead to attempts to place food in our mouths with our hands 2 feet over our heads—an ecological disaster! The distance effect is also useful at the other extreme, allowing us to reach accurately toward objects that are perceptually mislocalized.

THE GREAT CIRCLE MODEL FOR THE VISUAL INFLUENCE ON VPEL

Employing a projection sphere shifts the intellectual focus from the external stimulus to stimulation at the back of the eye and permits a broader generalization of the use of "retinal orientation." In this section we describe a geometric treatment that maps the distal stimulus onto the retina, the *Great Circle Model,* designed to answer questions about how the influences of individual inducer lines are organized (Matin & Li, 1992a, 1992b, 1994a, 1994c, 2001). Figure II.25-5 offers some insight regarding the way in which different retinal orientations combine their individual influences. Each panel contains a sketch of two individual lines in central projection on the inner surface of a sphere centered at the nodal point of the viewing eye in primary position. We refer to the projection sphere as a spherical approximation to the eye (SAE1) and the inner surface as its "retina" but with no intent of approximating the actual shape of the eyeball or its complexities (e.g., SAE1 assumes that the optical nodal point coincides with SAE1's center). We have recently developed an approximation, SAE2, in which the nodal point is 5 mm in front of the SAE2's center—approximating the actual distance in the emmetropic eye (Matin & Li, in preparation); the regularities and symmetries in SAE2 are much more complex than those we show in Figure II.25-4 for SAE1.

The use of spherical projection derives from the hypothesis that the identity of influence on VPEL by mirror

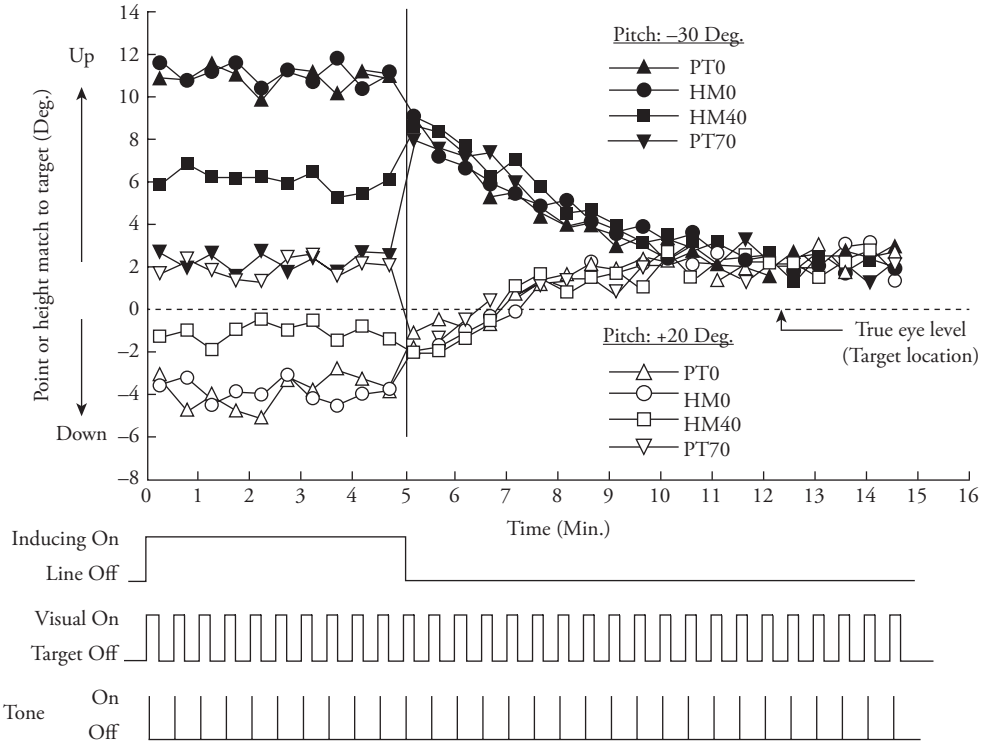


Figure II.25-4. The hand-to-body distance effect. Manual behavior during a 5-min period with the inducer on followed by further measurements for 10 min after inducer offset. Two values of inducer pitch (–30° or +20°) were used, with measurements for various combinations of three distances (0 cm, 40 cm, and 70 cm) and two manual behaviors—pointing (PT) and height matching (HM)—setting the height of the arm to match the perceived elevation of an eye level target. The behavior was measured every 30 s, and the results were averaged across four subjects. (Data from Li, Matin & Matin, 2013.)

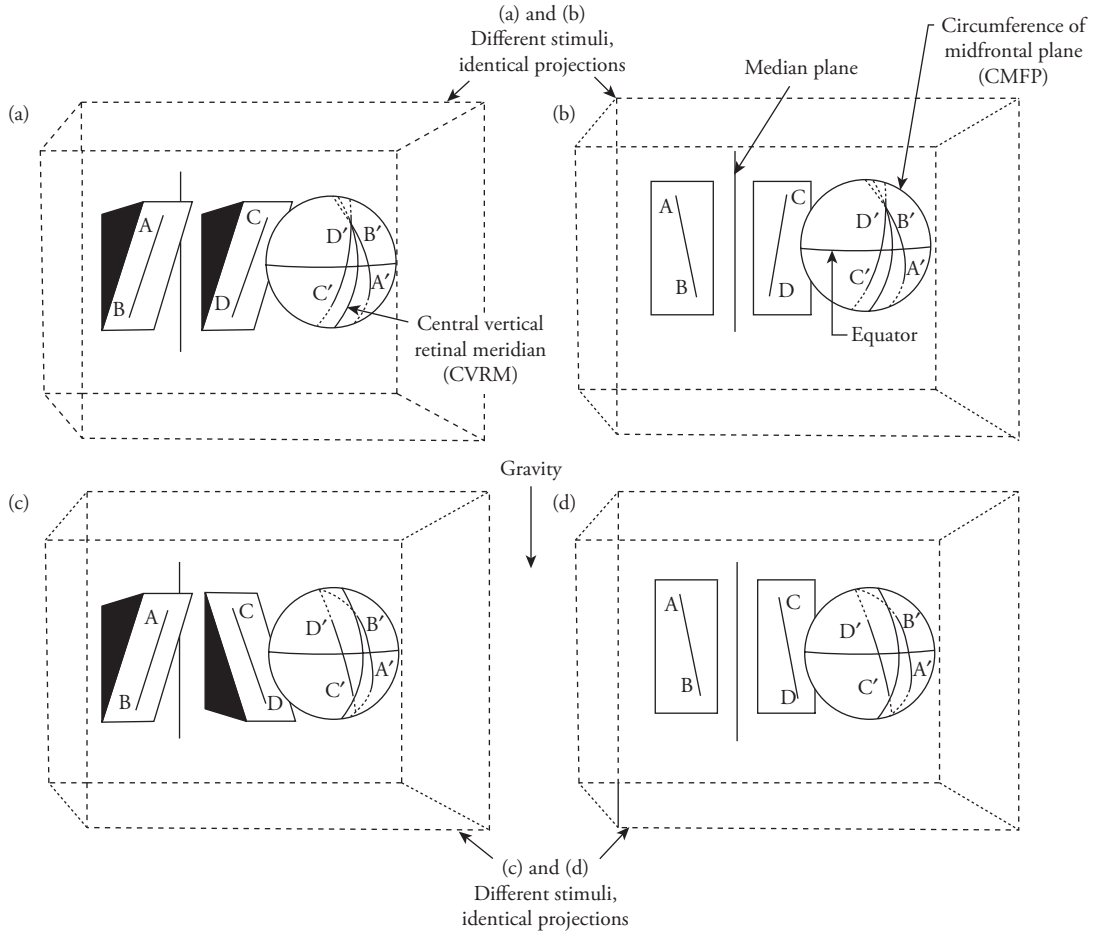


Figure II.25-5. Great circle model. Projection spheres centered on the eye of the erect observer in primary viewing position viewing a 2-line visual field. The central CVRM is the great circle that corresponds to the median plane of the eye of the observer; the CMFP is a frontoparallel section through the sphere; the equator is horizontal at the observer's eye level. In each of the four panels the two lines are at equal horizontal eccentricities on opposite sides of the median plane. Each of the two pairs of panels—(a) / (b), (c) / (d)—displays the identity of central projections on the projection sphere from two pitched-from-vertical lines on pitched-only plane(s) with two oblique lines on an erect frontoparallel plane. The two lines in (a) are parallel and lie in a single pitched-only plane whereas the two lines of "equivalent pitch" in its erect-plane counterpart in (b) are mirror symmetrical; the two lines in (c) lie in pitched-only planes of equal and opposite pitch whereas the two lines in its erect plane counterpart in (d) are parallel. Systematically changing the pitch of the plane containing a line from topbackward to topforward produces a linearly related change in the intersection of the great circle projection of the line with the CVRM. (Adapted from Matin & Li, 1999.)

symmetrical lines is a consequence of action by them at a common neural node. This connection builds on the fact that the image of every straight line in object space is imaged on a segment of a great circle on such a projection sphere and that the images of all members of a set of parallel lines fall on great circles that intersect each other at two nodes unique to the set; the nodes are different for different sets of parallel lines differentiated by their different directions in space. The two nodes for each set represent the point at infinity that lies on the extensions of the line images in the projection plane. As shown in Figure II.25-5a, all members of the set of parallel pitched-from-vertical lines on a pitched-only plane are imaged on segments of great circles that intersect at the same two points on the central vertical retinal meridian (CVRM; only one of these points is visible in each panel). The locations of these intersection points change linearly with pitch and map directly onto the variation in VPEL. Our hypothesis is that they also map topologically onto the neural nodes responsible for the visual influence on VPEL. These neural nodes are a prominent feature of the neuromathematical model that we describe in the next section.

In the present context, specification of retinal orientation for straight lines is most simply accomplished by designating the great circle on which the image falls. Although unique placement of a great circle on the two-dimensional spherical surface requires only two coordinates, it is useful to employ coordinates of intersection on three base great circles: the CVRM, the circumference of the midfrontal plane (CMFP), and the equator (Fig. II.25-5). This is because each of the three coordinates corresponds exactly to a conceptual variable of considerable interest that can be readily manipulated experimentally. For VPEL, the coordinate of most interest here is the point of intersection of the great circle corresponding to the image of a line with the CVRM; for pitched-from-vertical lines (and for lines of "equivalent pitch" from other lines imaged at the same retinal orientation, Fig. II.25-2a), this coordinate changes linearly with pitch. The point of intersection with the equator is a measure of the horizontal eccentricity of the line at

true eye level. The intersection of a line's great circle with the CMFP is a measure of the obliquity of the orientation within a frontal plane of an oblique line and its retinally identical pitched-from-vertical line. The intersections on the three base great circles are related by

$$\tan \mu_i = \tan \mu_j / \tan \mu_k$$

where μ_i, μ_j, and μ_k are the intersections on the CVRM, CMFP, and equator, respectively.

Since the focus of the Great Circle Model is on image location on the projection sphere, it readily accommodates identical results for lines on a pitched plane and tilted lines on an erect plane that strike the same retinal location (Figures II.25-2a, II.25-b). Importantly, the exponential increase with line length of the influence of the pitched line on the slope of the VPEL versus pitch function is the same whether the total length of the simultaneously exposed segment(s) are distributed equally between two symmetrically-placed lines or are coextensive (Fig. II.25-3a; Matin & Li, 1994b, 1999, 2000). Indistinguishable VPELs are generated by lines with the same intersection point on the CVRM but with opposite signs on the CMFP. These opposite retinal orientations are stimulated from mirror symmetric locations across the median plane either by oppositely oriented oblique lines on a frontal plane or by parallel lines on a pitched-only plane.

NEUROMATHEMATICAL TREATMENT OF THE VISUAL INFLUENCE ON VPEL

We have employed a three-stage neuromathematical model (Fig. II.25-6) for dealing with the influence of the inducer on VPEL (Matin & Li, 2001). Stage 1 parallel processes the orientations of the lines utilizing two classes of orientation-sensitive neural units in each hemisphere of visual cortex. The other two aspects of this model have been briefly described previously as the *weighted average treatment* (stage 3) and the *Great Circle Model* (stage 2). The model's output was well-fitted to the results of five sets of experiments with 1-line or 2-line inducing stimuli varying in orientation and length and to the entire well-illuminated, well-structured pitchroom.

THE BODY-REFERENCED MECHANISM

Felt Direction of Gaze and Gravity

Felt Direction of Gaze

In addition to its effects on VPEL and on felt arm level, the inducer affects the felt direction of gaze when observers are instructed to direct gaze elevation in a horizontal direction without a visual target. Two studies of eye

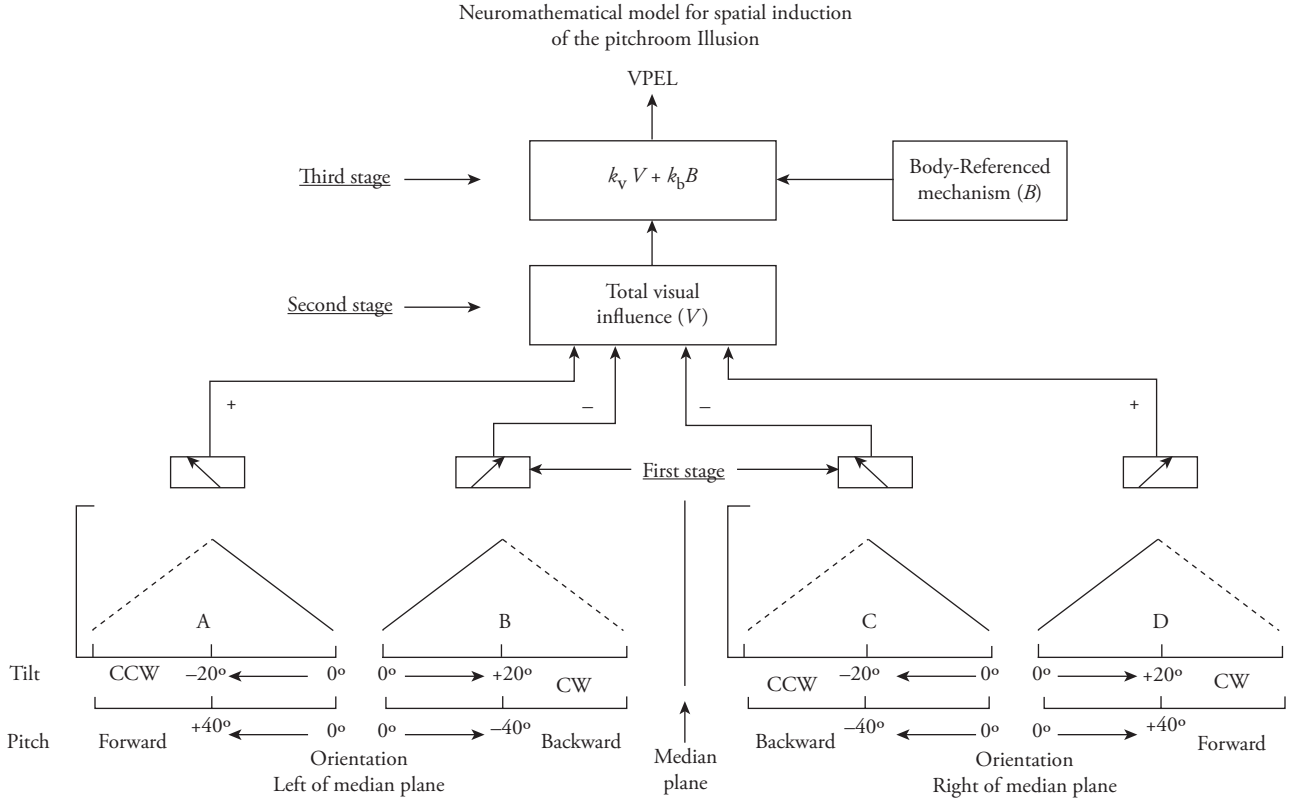


Figure II.25-6. Flow diagram of the three-stage neuromathematical model for VPEL. The first stage contains two classes of orientation-sensitive units in V1 on each side of the median plane; the second stage combines signals from all four classes of stage 1 units (see Fig. II.25-5 for the way in which images of lines of different lengths and orientations combine on a projection sphere); the third stage combines the total visual influence from stage 2 with influence from the body-referenced mechanism in a weighted average to generate VPEL. (Adapted from Matin & Li, 2001.)

position with such instructions showed vertical eye position versus inducer pitch functions that were remarkably similar to the VPEL versus inducer pitch function (Cohen, Ebenholtz, & Linder, 1995; Matin & Li, 1995b). This is the result that might be expected if the horizontal eye under induction with a topbackward stimulus "feels" that it is looking up and that the observer must therefore rotate it downward to feel horizontal—just as the observer in VPEL measurements with a topbackward inducer "sees" an eye-level target above eye level and sets a target below eye level to appear at eye level. A similar analysis applies, of course, with a topforward inducer (the eye "feels" as if it is looking down, so it looks up to "feel" horizontal).

The Influence of Gravity and the Elevator Illusion

The earliest experimental work on VPEL was done in human centrifuges in studies of the vestibular system in which the magnitude of the gravitoinertial vector was systematically varied without change in direction (Cohen, 1973, 1981; Correia, Hixson, & Niven, 1968). In these experiments, subjects set the elevation of a visual target to VPEL in otherwise total darkness. Increase in magnitude along the head-to-foot axis results in the elevator illusion in which a stationary visible target appears to rise. Other aspects of gravity's influence appeared in early reports on mislocalization in centrifuges (Graybiel, 1952; Whiteside, Graybiel, & Niven, 1965) and in manned orbital flight (Schöne, 1964), and later work on parabolic flight in airplanes (Cohen, 1992) added an important aspect to this work. Cohen's (1973) centrifuge experiments showed that in addition to the shear in the otolith organ (macular surface of the utricle), neck proprioception was also a factor in the elevator illusion.

More recent work in a human centrifuge discovered that the slope of the VPEL versus pitch function for a 2-line inducer decreased systematically with magnitude of the gravitoinertial vector from 1 g to 4 g and that the *y*-intercept—a measure of the magnitude of the elevator illusion—was lowered to about 19° below true eye level for the erect subjects (Chelette et al., 1995; see Li et al., 2001, for the specific results; also see Cohen, Stoper, Welch, & DeRoshia, 2001; Raphel & Barraud, 1994).

THE PITCHROOM ILLUSION WITHIN THE LARGE PICTURE

The pitchroom illusion, like other illusions, calls attention to psychological and neural processes that are obscured in the ordinary day-to-day interactions of people with their environments. The startling effects of pitched inducers on the perception of our bodies and of the outside world are among the facts that will shape a scientific picture of the multisensory representations of space and of our bodies within it. Moreover, as has been the case throughout the history of science, discoveries from research motivated by scientific curiosity also have promise for real-world applications. In the case of the pitchroom studies, we note two: the findings are relevant in the preparation of astronauts for space travel, and they are relevant in clinical work with patients who have suffered brain damage.

ACKNOWLEDGMENTS

This research was supported by AFOSR Grants 91-0146 and F49620-94-1-0397 from the US Air Force Office of Scientific Research, Grant EY10534 from the National Eye Institute of the National Institutes of Health, and Grant BCS-06-16654 from the National Science Foundation.

NOTES

1. The 1-line and 2-line inducers that influence VPEL also have large systematic influences on the classic rod-and-frame illusion in which a roll-tilted square frame induces a shift of the orientation in the frontoparallel plane corresponding to visually perceived vertical (VPV; Witkin, 1949; Witkin & Asch, 1948). VPV and VPEL follow identical rules of combination with one very important exception (Li & Matin, 2005b, 2005c; Matin & Li, 1994a): the VPEL versus pitch functions for inducing lines on the left and right of the median plane (Fig. II.25-2b) are not parallel to each other as they would be if the same retinal orientations on opposite sides of the observer's midline possessed identical influences. Instead, the variation of VPEL with retinal orientation follows bilaterally symmetric tracks. Thus VPEL increases with retinal orientation for the right line but decreases symmetrically with retinal orientation for the left line. However, the VPV versus retinal orientation functions for inducing lines on the left and right of the median plane are parallel to each other; equal and opposite retinal orientations have identical effects. Significant correlations were found for 30 observers between the induction effects on both VPEL and VPV that imply the existence of a single, shared visual orientation process that services separate integration processes for VPEL and VPV (Shavit, Li, & Matin, 2013). Of further interest to this comparison, a set of experiments employing functional magnetic resonance imaging showed that different cortical regions mediate the processing of a pair of bilaterally symmetric lines on bilaterally symmetric locations in the visual field (areas 40 and 7) from those that process parallel lines on the bilaterally symmetric locations (cortical area 19; Matin, Li, Hudson, & Hirsch, 2001). These results raise the possibility that different regions process the influence of the inducers on VPEL and VPV.

   Although nearly all of the work described in this chapter was done with monocular viewing, and, as noted, binocular viewing has been shown to have little if any influence on VPEL, binocular cues have powerful influences on many aspects of space perception, and horizontal binocular retinal disparity has been the centerpiece for understanding many unique aspects of stereovision. Particularly relevant in the present context centered on changes in orientation resulting from induction by lines is the well-known and much-studied binocular depth contrast illusion, the induction of an appearance of visual pitch in a centrally located erect line by pitched-from-vertical flanking lines (Ogle, 1946, 1950; Werner, 1937, 1938), which, as has recently been shown, does not cross the median plane but is restricted to the side of the visual field containing a unilateral inducer (Hudson, Matin, & Li, 2008). These results cannot be explained by retinal disparity alone or in combination with any additional factors or processes previously suggested to account for stereovision. However, a two-channel, three-stage neuromathematical model accounts for those and previous results along with a number of the most prominent features of binocular pitch perception (Hudson et al., 2008).

REFERENCES

Chelette, T., Li, W., Esken, R., & Matin, L. (1995). Visual perception of eye level (VPEL) under high *g* while viewing a pitched visual field. Paper presented at the annual meeting of Aerospace Medical Association. Chicago, May.

Cohen, M. M. (1973). Elevator illusion: Influences of otolith organ activity and neck propioception. *Perception & Psychophysics*, *14*, 401–406.

Cohen, M. M. (1981). Visual-proprioceptive interactions. In R. D. Walk & H. L. Pick Jr. (Eds.), *Intersensory perception and sensory integration* (pp. 175–215). New York: Plenum Press.

Cohen, M. M. (1992). Perception and action in altered gravity. *Annals of the New York Academy of Sciences*, *656*, 354–362.

Cohen, M. M., Ebenholtz, S. M., & Linder, B. J. (1995). Effects of optical pitch on oculomotor control and the perception of target elevation. *Perception & Psychophysics*, *57*, 433–440.

Cohen, M. M., Stoper, A. E., Welch, R. B., & DeRoshia, C. W. (2001). Effects of gravitational and optical stimulation on the perception of target elevation. *Perception & Psychophysics*, *63*, 29–35.

Correia, M. J., Hixson, W. C., & Niven, J. J. (1968). On predictive equations for subjective judgments of vertical and horizon in a force field. *Acta Otolaryngologica*, *230*(Suppl.), 3–20.

DiZio, P., Li, W., Lackner, J. R., & Matin, L. (1997). Combined influences of gravitoinertial force level and visual field pitch on visually perceived eye level. *Journal of Vestibular Research*, *7*, 381–392.

Graybiel, A. (1952). The oculogravic illusion. *AMA Archives of Ophthalmology*, *48*, 605–615.

Hudson, E. T., Li, W., & Matin, L. (2000). Independent mechanisms produce visually perceived eye level (VPEL) and perceived visual pitch (PVP). *Vision Research*, *40*, 2605–2619.

Hudson, E. T., Li, W., & Matin, L. (2006). The field dependence/independence cognitive style does not control the spatial perception of elevation. *Perception & Psychophysics*, *68*, 377–392.

Hudson, E. T., Li, W., & Matin, L. (2008). Binocular spatial induction for the perception of depth does not cross the midline. *Proceedings of the National Academy of Sciences USA*, *105*, 18006–18011.

Li, W., Dallal, N., & Matin, L. (2001). Influences of visual pitch and visual yaw on visually perceived eye level (VPEL) and straight ahead (VPSA) for erect and rolled-to-horizontal observers. *Vision Research*, *41*, 2873–2894.

Li, W., & Matin, L. (1992). Rotation in depth of linear arrays of points systematically influences egocentric localization. *Bulletin of the Psychonomic Society*, *30*, 453.

Li, W., & Matin, L. (1996). Visually perceived eye level is influenced identically by lines from erect and pitched planes. *Perception*, *25*, 831–852.

Li, W., & Matin, L. (1998). Change in visually perceived eye level without change in perceived pitch. *Perception*, *27*, 553–572.

Li, W., & Matin, L. (2005a). Two wrongs make a right: Linear increase of accuracy of visually guided manual pointing, reaching, and height-matching with increase in hand-to-body distance. *Vision Research*, *45*, 533–550.

Li, W., & Matin, L. (2005b). The rod-and-frame effect: The whole is less than the sum of its parts. *Perception*, *34*, 699–716.

Li, W., & Matin, L. (2005c). Visually perceived vertical (VPV): Induced changes in orientation by 1-line and 2-line roll-tilted and pitched visual fields. *Vision Research*, *45*, 2037–2057.

Li, W., Matin, E., Bertz, J. W., & Matin, L. (2008). A tilted frame deceives the eye *and* the hand. *Journal of Vision*, *8*(16), 18.1–18.16.

Li, W., Matin, E., & Matin, L. (2013). Short-lived effects of a visual inducer during egocentric space perception and manual behavior. *Attention, Perception &Psychophysics*, *75*, 1012–1026.

MacDougall, R. (1903). The subjective horizon. *Psychological Monographs*, *4*(1, Suppl. 4), 145–166.

Matin, L., & Fox, C. R. (1986). Perceived eye level: elevation jointly determined by visual field pitch, EEPI, and gravity. *Investigative Ophthalmology & Visual Science*, *27*(Suppl.), 333.

Matin, L., & Fox, C. R. (1989). Visually perceived eye level and perceived elevation of objects: Linearly additive influences from visual field pitch and from gravity. *Vision Research*, *29*, 315–324.

Matin, L., & Fox, C. R. (1990). Visually perceived eye level and perceived elevation of objects: linearly additive influences from visual field pitch and from gravity. Corrigenda. *Vision Research*, *30*(1), I.

Matin, L., & Li, W. (1992a). Mislocalizations of visual elevation and visual vertical induced by visual pitch: The Great Circle Model. *Annals of the New York Academy of Sciences*, *656*, 242–265.

Matin, L., & Li, W. (1992b). Visually perceived eye level: Changes induced by a pitched-from-vertical 2-line visual field. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 257–289.

Matin, L., & Li, W. (1994a). Mirror symmetry and parallelism: Two opposite rules for the identity transform in space perception and their unified treatment by the Great Circle Model. *Spatial Vision*, *8*, 469–489.

Matin, L., & Li, W. (1994b). Spatial summation among parallel lines across wide separation (50°): Spatial localization and the Great Circle Model. *Vision Research*, *34*, 2577–2598.

Matin, L., & Li, W. (1994c). The influence of the orientation of a stationary single line in darkness on the visual perception of eye level. *Vision Research*, *34*, 311–330.

Matin, L., & Li, W. (1995a). Light and dark adaptation of visually perceived eye level controlled by visual pitch. *Perception & Psychophysics*, *57*, 84–104.

Matin, L., & Li, W. (1995b). Multimodal basis for egocentric spatial localization and orientation. *Journal of Vestibular Research*, *5*, 499–518.

Matin, L., & Li, W. (1999). Averaging and summation of influences on visually perceived eye level between two long lines differing in pitch or roll-tilt. *Vision Research*, *39*, 307–329.

Matin, L., & Li, W. (2000). Linear combinations of signals from two lines of the same or different orientations. *Vision Research*, *40*, 517–527.

Matin, L., & Li, W. (2001). Neural model for processing the influence of visual orientation on visually perceived eye level (VPEL). *Vision Research*, *41*, 2845–2872.

Matin, L., & Li, W. (2010). Combined influences of extraretinal signals, retinal signals, and visual induction on space perception and manual behavior in perisaccadic and steady viewing. In R. Nijhawan & B. Khurana (Eds.), *Problems of space and time in perception and action* (pp. 63–93). Cambridge, UK: Cambridge University Press.

Matin L., & Li, W. (in preparation). A new model for the projection sphere as a spherical approximation to the eye.

Matin, L., Li, W., Hudson, T., & Hirsch, J. (2001). Perceptions of elevation and orientation: From the stimulus basis to the cerebral cortex. *Journal of Vision*, *1*, 56.

Matin, L., Li, W., Li, L., & Shavit, A. Y. (2006). Influence of roll-tilt, interpoint separation, and length of linear points-arrays on a frontoparallel plane on visually perceived eye level (VPEL). *Journal of Vision*, *6*, 972.

Matin, L., Matin, E., & Li, W. (2016). Dipole analysis of the influence of linear array of points on visually perceived eye level (VPEL). *Vision Research*, *126*, 254–263.

Matin, L., Picoult, E., Stevens, J. K., Edwards, M. W. Jr., Young, D., & MacArthur, R. (1982). Oculoparalytic illusion: Visual-field dependent mislocalizations by humans partially paralyzed with curare. *Science*, *216*, 198–201.

Matin, L., Stevens, J. K., & Picoult, E. (1983). Perceptual consequences of experimental extraocular muscle paralysis. In A. Hein & M. Jeannerod (Eds.), *Spatially oriented behavior* (pp.243–262). New York: Springer.

Ogle, K. N. (1946). The binocular depth contrast phenomenon. *American Journal of Psychology*, *59*, 111–126.

Ogle, K. N. (1950). *Researches in binocular vision*. New York: Hafner.

O'Shea, R. P., & Ross, H. E. (2007). Judgments of visually perceived eye level (VPEL) in outdoor scenes: Effects of slope and height. *Perception*, *36*, 1168–1178.

Poquin, D., Ohlmann, T., & Barraud, P. A. (1998). Isotropic visual field effect on spatial orientation and egocentric localization. *Spatial Vision*, *11*, 261–278.

Post, R. B., Teague, J. M., Welch, R. B., & Hudson, T. E. (2003). Visually perceived eye level with reversible pitch stimuli: Implications for the great circle and implicit surface models. *Perception*, *32*, 1073–1092.

Post, R. B., & Welch, R. B. (1996). The role of retinal versus perceived size in the effects of pitched displays on visually perceived eye level. *Perception*, *25*, 853–859.

Post, R. B., Welch, R. B., & Clark, V. D. (2000). Why do pitched horizontal lines have such a small effect on visually perceived eye level? *Perception & Psychophysics*, *62*, 726–734.

Raphel, C., & Barraud, P. A. (1994). Perceptual thresholds of radial accelerations as indicated by visually perceived eye level. *Aviation, Space, Environmental Medicine, 65*, 204–208.

Robison, E., Li, W., & Matin, L. (1995). Manual matches to the elevation of visually mislocalized targets. *Investigative Ophthalmology & Visual Science 36*(Suppl.4), S358.

Rossetti, Y. (1998). Implicit perception in action: Short-lived motor representations of space. *Consciousness and Cognition*, *7*, 520–558.

Schöne, H. (1964). On the role of gravity in human spatial orientation. *Aerospace Medicine*, *35*, 764–772.

Servos, P., Matin, L., & Goodale, M. A. (1995). Dissociation between two modes of spatial processing by a visual form agnosic. *NeuroReport*, *6*, 1893–1896.

Shavit, A. Y., Li, W., & Matin, L. (2004). Influences of global and local orientations of line segments on perceived eye level. Paper presented at the 45th annual meeting of the Psychonomic Society. Minneapolis, November.

Shavit, A. Y., Li, W., & Matin, L. (2008). The influences of array orientation and of line orientation on visually perceived eye level (VPEL) are modulated by line length and array length. *Journal of Vision*, *8*(6), 457.

Shavit, A. Y., Li, W., & Matin, L. (2013). Individual differences in perceived elevation and verticality: Evidence of a common visual process. *Multisensory Research*, *26*, 205–239.

Stevens, J. K., Emerson, R. C., Gerstein, G. L., Kallos, T., Neufeld, G. R., Nichols, C. W., & Rosenquist, A. C. (1976). Paralysis of the awake human: Visual perceptions. *Vision Research*, *16*, 93–98.

Stoper, A. E. (1998). Height and extent: Two kinds of perception. In E. Winograd, R. Fivush, & W. Hirst (Eds.), *Ecological approaches to cognition: Essays in honor of Ulric Neisser* (pp. 97–124). Hillsdale, NJ: Lawrence Erlbaum.

Stoper, A. E., & Cohen, M. M. (1989). Effect of structured visual environments on apparent eye level. *Perception & Psychophysics*, *46*, 469–475.

Welch, R. B., & Post, R. B. (1996). Accuracy and adaptation of reaching and pointing in pitched visual environments. *Perception & Psychophysics*, *58*, 383–389.

Werner, H. (1937). Dynamics of binocular depth perception. *Psychological Monographs*, *49*(2), 1–121.

Werner, H. (1938). Binocular depth contrast and the conditions of the binocular field. *American Journal of Psychology*, *51*, 489–497.

Whiteside, T. C. D., Graybiel, A., & Niven, J. I. (1965). Visual illusions of movement. *Brain*, *88*, 193–210.

Witkin, H. A. (1949). Perception of body position and of the position of the visual field. *Psychological Monograph*, *63*(7), 1–46.

Witkin, H. A., & Asch, S. E. (1948). Studies in space perception: IV. Further experiments on perception of the upright with displaced visual fields. *Journal of Experimental Psychology*, *38*, 762–782.

Chapter 26
Geometric Illusions in the Human Face and Body

Kazunori Morikawa

INTRODUCTION

Studies of visual geometric illusions began more than a century ago with simple geometric illusions of orientation and length, such as the Zöllner illusion and the Müller-Lyer illusion. Historically, most geometric illusions have been studied with very little consideration for practical applications. Some visual illusions, however, can be very useful in daily life. For example, clothing and cosmetic makeup take advantage of visual illusions so as to make the human body and face look more attractive. Fashion stylists have the expertise to make people look slimmer and make their legs look longer (e.g., Feldon, 2000; Yokomori, 2007). Makeup artists have standard conventions to make eyes look larger and cheeks look narrower. These are the geometric illusions in natural objects that we encounter every day.

Except for few studies, such as Morikawa (2003) on the Müller-Lyer illusion and Thompson and Mikellidou (2011) and Ashida, Kuraguchi, and Miyoshi (2013) on the Helmholtz illusion, these real-life illusions have been largely overlooked by vision scientists and have seldom been measured or investigated psychophysically. Therefore, it is of interest to study these illusions for both scientific value and practical usefulness. This chapter reviews several classes of these bodily illusions.

THE MÜLLER-LYER ILLUSION IN CLOTHING

Morikawa (2003) showed that a high-cut bathing suit makes legs look longer by approximately 5%. Using correlation coefficients, he also indicated that there are common mechanisms underlying the high-cut leg illusion and the Müller-Lyer illusion (Fig. II.26-1).

THE DELBOEUF ILLUSION IN EYE MAKEUP

It is well known in the cosmetics industry that an application of eyeliner makes eyes look larger (Fig. II.26-2a). This illusion is most likely based on the Delboeuf illusion, in which the inner circle becomes assimilated to the outer circle and looks larger (Morikawa et al., 2015). Likewise, double-edged eyelids increase the apparent size of eyes because of the Delboeuf illusion (Morikawa, 2012). Perceived eye size can also be influenced by other geometric illusions, some of which act in opposing directions (Morikawa, 2012).

There exist very few studies that have psychophysically measured the amount of illusion caused by makeup. Morikawa and Fujii's (2009) experiments used the method of adjustment to measure the perceived size of eyes with makeup. The stimuli were computer-generated faces with either standard size eyes or 15% enlarged eyes. Eyeliner and/or eye shadow was applied to those faces (Fig. II.26-2b). The comparison stimuli were contours made by tracing the right eye of the face, which were then enlarged or reduced in size so as to create a series of the same shape in different sizes. Participants made the size of a comparison stimulus larger or smaller until it matched the perceived size of the eye. Figure II.26-2c shows part of the results. The application of eyeliner resulted in a maximum of 7.3% overestimation (or 15% in area) of eye size for the standard size eye. However, for the 15% enlarged eye, eyeliner caused very little effect, suggesting that there is a natural upper limit to the magnitude of geometric illusions in the human face. Recently, we psychophysically measured the illusory effects of five levels of eyeliner combined with five levels of mascara using the staircase method (Matsushita, Morikawa, & Yamanami, 2015). We found that mascara outweighed eyeliner in making eyes appear larger.

THE BICOLOR ILLUSION IN CLOTHING

There are several ways to make a person look slimmer than he or she actually is, such as the bicolor illusion (Morikawa, 2012; Tezuka, Uemura, & Morikawa, 2012; see Fig. II.26-3, top panel). The bicolor illusion probably arises from two factors. First, a small number of divisions (one or two) within a rectangle cause a perceptual shrinkage of the rectangle along the divided dimension, although a higher number of divisions cause a perceptual expansion, such as the Oppel-Kundt illusion (Pinna, 2012). Second, in the figure, each strip in the tripartite rectangle on the right is long and narrow, which accentuates the rectangle's vertical dimension. The bottom panel of Figure II.26-3 shows one of the fashion applications of the bicolor illusion (Feldon, 2000). The long vest makes the person look slimmer and taller by drawing viewer's attention to her vertical dimension and dividing her torso width into three parts.

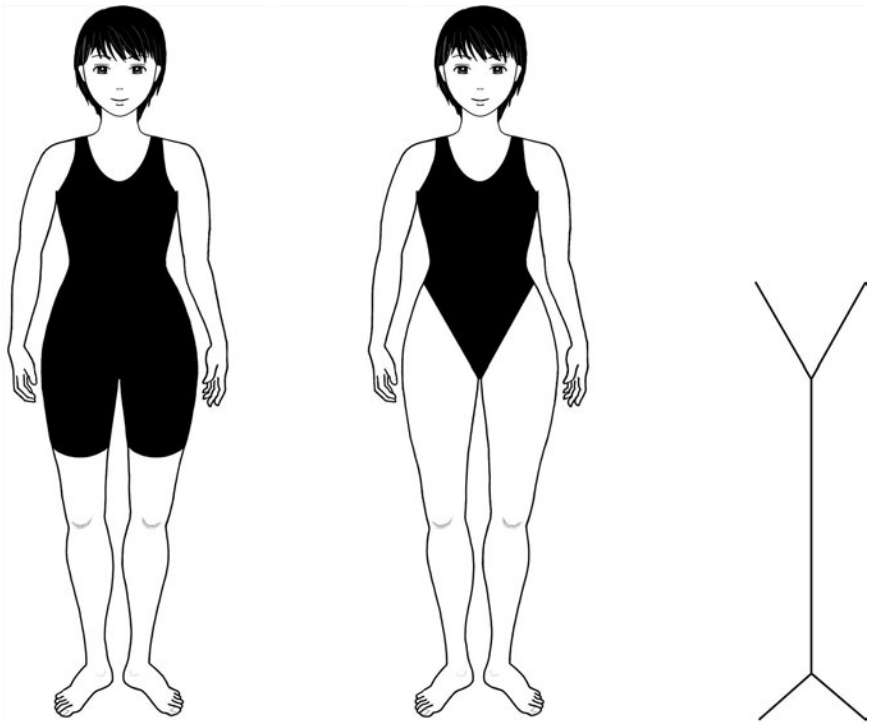


Figure II.26-1. High-cut leg illusion. The woman on the right appears to have longer legs than the woman on the left. This is an application of the Müller-Lyer illusion on the right. (From Morikawa, 2012).

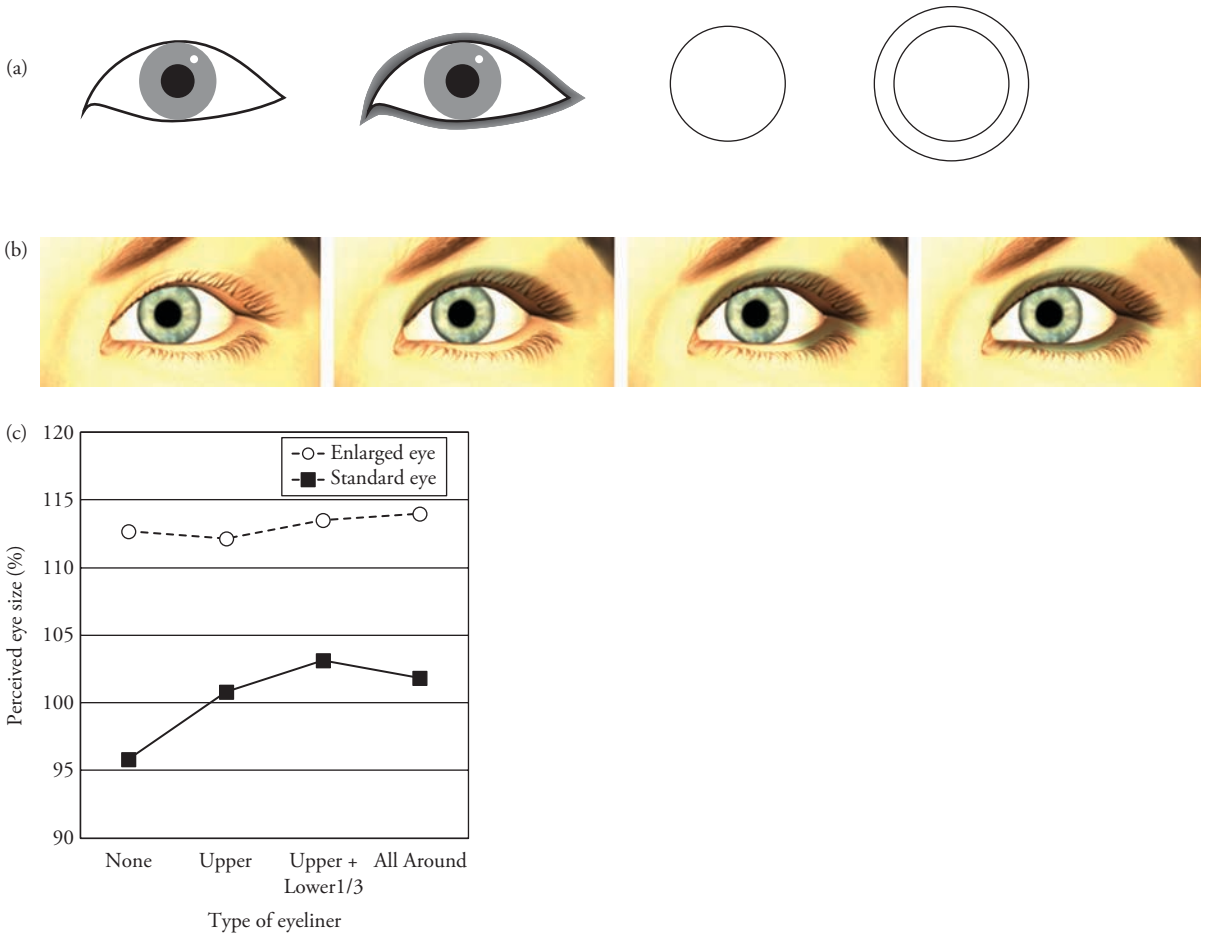


Figure II.26-2. Eyeliner as an application of the Delboeuf illusion. (a) An eye with eyeliner (second from left) appears larger than the same eye without eyeliner (leftmost), in the same way as the inner circle (rightmost) appears larger than a single circle of the same size (second from right), which is the Delboeuf illusion. (b) Eye part of the facial stimuli used in the experiments of Morikawa and Fujii (2009). From left to right: No eyeliner, upper eyeliner, upper plus lower one-third eyeliner, and eyeliner all around the eye. (c) Part of the results of Morikawa and Fujii (2009) showing perceived eye size as a function of types of eyeliner. (Adapted from Morikawa, 2012.)



Figure II.26-3. Top panel: In the bicolor illusion, the tripartite rectangle on the right appears slimmer and taller than the bipartite rectangle on the left, which appears stockier. Bottom panel: The woman on the right wearing a long vest appears slimmer than the woman on the left. (From Morikawa, 2012.)

SHAPE ECHO ILLUSIONS IN THE HUMAN FACE AND BODY

In Figure II.26-4a, the head size of the person on the right appears larger than that of the person on the left. Here the term "head size" refers to the width of the upper head above the eyes. In other words, the person on the right seems to need a larger-sized hat than the person on the left. This impression is especially pronounced if we look at the face rather than the head per se. In fact, their heads are exactly the same size, because the two images are identical except for the width of cheeks, jaws, and necks. This is called the head size illusion (Morikawa, Okumura, & Matsushita, 2012). Using psychophysical experiments, Morikawa et al. (2012) found that fatter lower faces caused the head size to be 4% overestimated and thinner lower faces caused the head size to be 3% underestimated. They also showed that approximately 50% of the illusion magnitude results from holistic (configural) processing of faces. It should be noted that the head size illusion is very different from the face illusion reported by Lee and Freire (1999) in which moving the eyes and mouth toward the nose causes an oval aperture windowing the face to be perceived as shorter and wider. In the head size illusion, the configuration of facial features is never changed. It is the lower-half outline shape of the face that causes the illusion in the upper-half outline shape.

The head size illusion is an example of a class of illusions in which, if a part of the face or body deviates from the average shape in one way (e.g., wider than average), it distorts the perception of adjacent parts in the same direction (i.e., wider than they really are). This phenomenon can be regarded as a shape property of one area spreading to adjacent areas in much the same way as echoes spread and are reverberated. Therefore, we refer to these illusions as "shape echo illusions."

Figure II.26-4b shows another example of shape echo illusions (Morikawa, 2012). The woman on the left appears

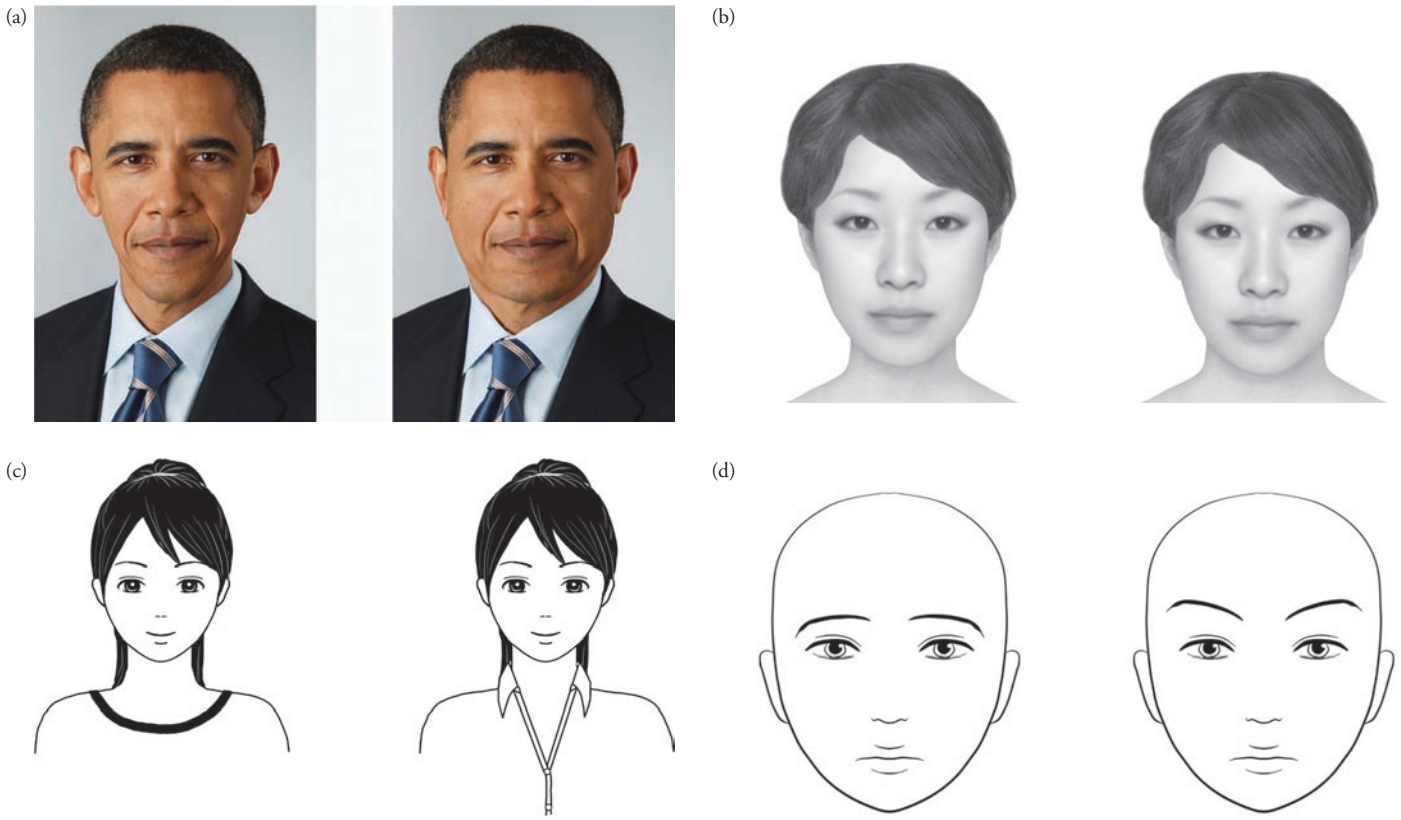


Figure II.26-4. Various "shape echo illusions." (a) The head size (or hat size) of the person on the right appears larger than that of the person on the left. However, both faces are identical except for the shape of the cheeks and jaws. (b) The woman on the left appears to have narrower jaws and cheeks than the woman on the right. However, the only difference is the shape of the neck. (c) The woman on the left appears to have wider, rounder jaws and cheeks than the woman on the right. However, the only difference is clothing. (d) The face on the left appears to have drooping eyes and rounder cheeks. The face on the right appears to have upward-slanted eyes and narrower, sharper cheeks. However, the only difference between the two faces is the eyebrows. (From Morikawa, 2012.)

to have narrower jaws and cheeks than the woman on the right. Actually, their faces are identical. The only difference is the shape of the neck. A narrow neck makes jaws and cheeks look narrower and sharper, whereas a wide neck makes them look wider and rounder. If "size contrast" occurred between the neck and the face, the narrow neck would make the face look wider. However, the reverse is actually perceived. It seems that between body parts "size assimilation" is more likely to occur than "size contrast." Figure II.26-4b suggests that if one wears a turtleneck sweater that makes one's neck look thicker, one's face will also look wider.

Figure II.26-4c shows an example of shape echo illusions caused by clothing (Morikawa, 2012). Wearing a crewneck sweater or a T-shirt makes one's face look rounder, whereas wearing a shirt with a sharp collar that creates a V shape makes one's face look narrower and sharper. Although I have yet to psychophysically measure the illusions shown in Figures II.26-4b and II.26-4c, it seems that some people see these illusions more clearly than do others. It is possible that some perceptual learning contributes to these facial illusions.

Furthermore, in facial perception the shape of one part of a face affects the perception of other parts. Makeup can take advantage of these illusions. In Figure II.26-4d, the only difference between the two faces is the eyebrows (Morikawa, 2012). Nevertheless, the drooping eyebrows make the eyes appear to be drooping and make the cheeks look rounder. Eyebrows that are slanted upward make the eyes also look slanted upward, thus making the cheeks look narrower and sharper. This means the shape and orientation of the eyebrows can influence the perception of a large portion of a face. Using psychophysical experiments, Matsushita et al. (2015) demonstrated that slanting the eyebrows by 16° caused a perceived slanting of the eyes by 3° in the same direction.

ILLUSIONS OF AMODAL PERCEPTUAL COMPLETION IN CLOTHING

In the real world, many objects we see are partially obscured by other objects in front of them. The human visual system is well adapted to recognize partially occluded objects. When we see a partially occluded object, our visual system automatically estimates what the occluded parts would look like. However, such amodal completion is not always accurate. Amodal completion is often biased toward the visible parts. For example, the woman on the left in the top panel of Figure II.26-5 appears to have slimmer legs than the woman on the right. Actually, the two women are identical except for the positions of the occluding bars. When only the

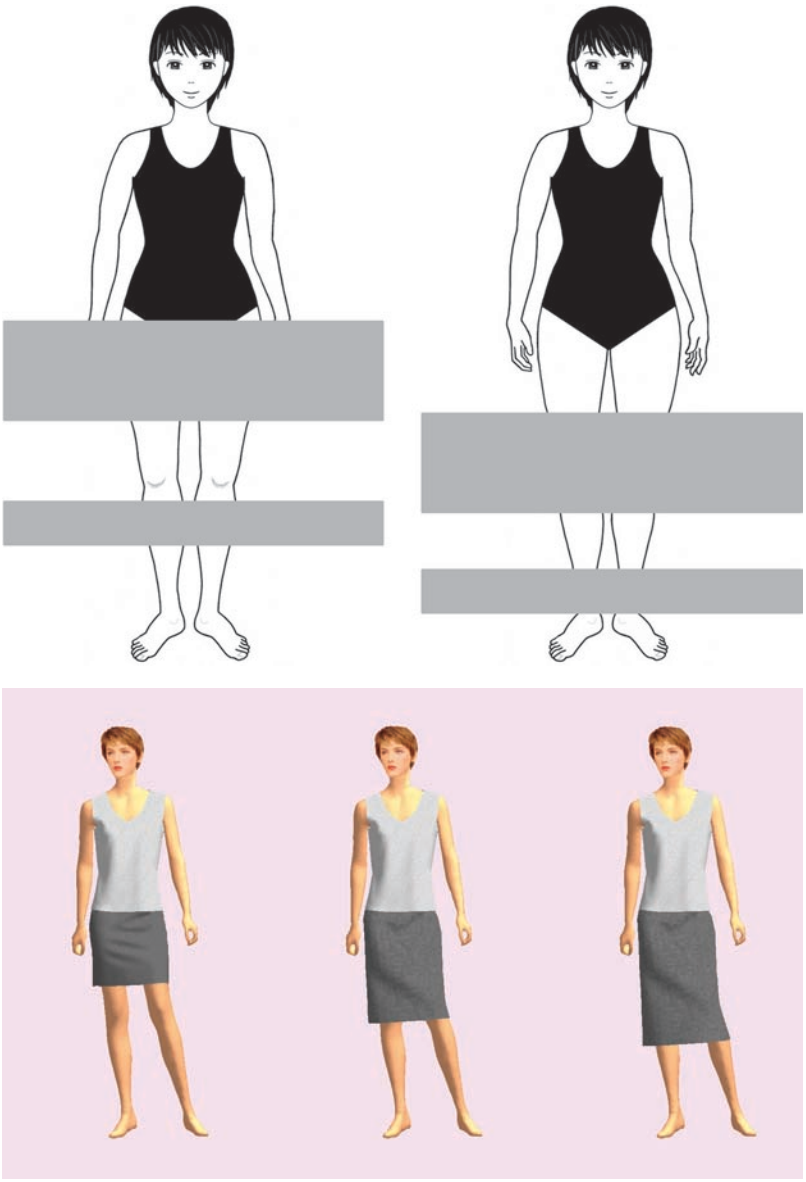


Figure II.26-5. Illusions of amodal perceptual completion. Top panel: The woman on the left appears to have slimmer legs than the woman on the right. Bottom panel: Legs look longest in the above-knee skirt (left) and look shortest in the below-knee skirt (right). (From Morikawa, 2012.)

thin parts of the legs, such as knees and ankles, are visible, the entire legs appear thin. Conversely, when only the thick parts of the legs, such as thighs and calves, are exposed, the entire legs appear thick. This phenomenon is often utilized in clothing. For example, three-quarter-length sleeves and cropped pants expose only the thinnest parts of the arms and legs, thus making the entire arms and legs appear thinner. In ordinary geometric illusions it is the visible parts whose length, shape, or size is misperceived, whereas in illusions of amodal perceptual completion, it is the invisible parts whose length, shape, or size is misestimated. Nevertheless, biased amodal completion causes perceived length, shape, or size of the human body to be different from reality. In this sense, illusions of amodal perceptual completion can be considered geometric illusions.

In addition, the perceived length of legs can be influenced by inaccurate amodal completion. Again, amodal completion is biased by visible parts. If the visible parts are long, the entire leg tends to appear long. In the bottom panel of Figure II.26-5, the above-knee skirt (left) makes legs look longer and the below-knee skirt (right) makes legs look shorter than does the knee-length skirt (middle). We conducted experiments using stimuli such as the bottom panel of Figure II.26-5 and the method of adjustment so as to measure the illusion amount of perceived leg length caused by skirts of various designs and lengths. The results consistently showed that above-knee skirts make legs look longer and below-knee skirts make legs look shorter than they actually are. The maximum amount of the leg length illusion was approximately 5% (Morikawa, 2012).

Inaccurate amodal completion can also lead to an illusion of facial shape. Okumura and Morikawa's (2010) experiments showed that hairstyles that cover facial outlines cause facial width to be perceived as narrower.

DISCUSSION

In contrast illusions such as the Ebbinghaus illusion, large objects make a neighboring small object appear even

smaller. However, in Figure II.26-4a, the wider lower-half shape of the face causes the upper-half shape to appear wider than it actually is, which can be considered an assimilation illusion. Likewise, the shape echo illusions in Figures II.26-4b, II.26-4c, and II.26-4d are all assimilation illusions. Moreover, light skin tone in the neck and décolleté area makes the adjacent facial skin tone appear lighter than it actually is (Futagawa, Hirayama, & Yamazaki, 2009), which is an illusion of assimilation and can be considered a lightness echo illusion. Futagawa et al. also found that, when the stimulus was perceived as a non-face object, the illusion of lightness contrast occurred. Therefore, illusions in the human face and body tend to arise in the direction of assimilation rather than in the direction of contrast.

One reason for the predominance of assimilation may be the fact that spaces between the lines that constitute classical geometric illusions are empty, whereas spaces between facial parts or between body parts are filled and connected with tissue such as skin, muscles, and bones. Therefore, if an individual's eyebrows are drooping, his or her eyes are likely to droop as well. If one part of the body is fat, other parts are likely to be fat as well. Perhaps, shape echo illusions arise because the visual system takes these natural co-occurrences and correlations into account.

When geometric illusions in the human face and body are psychophysically measured, the maximum illusion magnitude is typically 4% to 6% (Matsushita, Morikawa, & Yamanami, 2015; Morikawa, 2003, 2012; Morikawa & Fujii, 2009; Morikawa et al., 2012). People are often surprised at these small numbers, because changes in impression of the face and body caused by makeup and clothing seem striking (otherwise people would not buy cosmetics or certain clothes). This discrepancy between the psychophysical magnitude of bodily illusions and subjective impression probably arises from the fact that the human visual system is particularly tuned to discriminating fine differences in the human face and body. Identifying and recognizing faces is extremely important in society. We are so sensitive to the configuration of facial features that shifting eye position in a life-size facial photograph by even a few millimeters causes a large change in impression. Likewise, judging a person's body size and body proportions is important for social decision-making, such as liking or disliking, fight or flight, and dominating or obeying. Consequently, we are very sensitive to body size and body proportion. In other words, our difference threshold for the human face and body is very small. Therefore, even a change of a few percentage points becomes quite noticeable and makes a deep impression.

As described earlier, geometric illusions in the human face and body differ from classical geometric illusions consisting of simple lines in a few aspects. I propose that these bodily illusions be collectively called "biological illusion." In the study of motion perception, "biological motion" has been established as a distinctive field. Likewise, the study of biological illusion may become an exciting new and practical field of perceptual research in the future. It may bridge the gap between perception and impression, perception and aesthetics, perception and cosmetics, and perception and fashion.

REFERENCES

Ashida, H., Kuraguchi, K., & Miyoshi, K. (2013). Helmholtz illusion makes you look fit only when you are already fit, but not for everyone. *i-Perception*, *4*, 347–351.

Feldon, L. (2000). *Does this make me look fat?* New York: Villard.

Futagawa, A., Hirayama, K., & Yamazaki, K. (2009). The effect of skin color in décolleté on the impression of facial skin brightness. *Kaogaku (Journal of Japanese Academy of Facial Studies)*, *9*, 240. (in Japanese)

Lee, K., & Freire, A. (1999). Effects of configuration change on shape perception: A new illusion. *Perception*, *28*, 1217–1226.

Matsushita, S., Morikawa, K., Mitsuzane, S., & Yamanami, H. (2015). Eye shape illusions induced by eyebrow positions. *Perception*, *44*, 529–540.

Matsushita, S., Morikawa, K., & Yamanami, H. (2015). Measurement of eye size illusion caused by eyeliner, mascara, and eye shadow. *Journal of Cosmetic Science*, *66*, 161–174.

Morikawa, K. (2003). An application of the Müller-Lyer Illusion. *Perception*, *32*, 121–123.

Morikawa, K. (2012). New directions in research on visual illusions of shape and size related to the human face and body: Illusions caused by makeup and clothing. *Japanese Psychological Review*, *55*, 348–361. (in Japanese with an English abstract)

Morikawa, K., Matsushita, S., Tomita, A., & Yamanami, H. (2015). A real-life illusion of assimilation in the human face: eye size illusion caused by eyebrows and eye shadow. *Frontiers in Human Neuroscience*, *9*, 139. doi: 10.3389/fnhum.2015.00139

Morikawa, K., & Fujii, Y. (2009). Measuring the illusory effect of eye makeup. *Kaogaku (Journal of Japanese Academy of Facial Studies)*, *9*, 242. (in Japanese)

Morikawa, K., Okumura, K., & Matsushita, S. (2012). Head size illusion: Head outlines are processed holistically too. *Perception*, *41*(ECVP Abstract Supple.), 115.

Okumura, K., & Morikawa, K. (2010). Do hairstyles affect perception of face size? *Kaogaku (Journal of Japanese Academy of Facial Studies)*, *10*, 177. (in Japanese)

Pinna, B. (2012). The role of amodal completion in shape formation: Some new shape phenomena. *Perception*, *41*, 1336–1354.

Tezuka, H., Uemura, U., & Morikawa, K. (2012, February). Illusion magic in fashion and makeup. *CanCam*, pp. 211–218. Tokyo: Shogakukan Inc.

Thompson P., & Mikellidou K. (2011). Applying the Helmholtz illusion to fashion: Horizontal stripes won't make you look fatter. *i-Perception*, *2*, 69–76.

Yokomori, M. (2007). *Zukai kiyase magic [Magically slim-looking clothes illustrated]*. Tokyo: Shogakukan Inc.

Chapter 27
Dynamic Illusory Size Contrast

Enhanced Relative Size Effects Due to Stimulus Motion

Ryan E. B. Mruczek, Christopher, D. Blair, Lars Strother, and Gideon P. Caplovitz

INTRODUCTION

Our environment is highly dynamic. We move our eyes, heads, and bodies, and objects themselves may move of their own accord or through the influence of other animate forces. Although motion is a ubiquitous feature of our visual world, it has often been ignored in studies of size perception. For example, although classic size illusions such as the Delboeuf (Delboeuf, 1865; Nicolas, 1995) and Ebbinghaus (Burton, 2001; Thiéry, 1896) illusions (see chapter 28 in this volume) demonstrate how the relative size of a surrounding context can influence the perceived size of a target object, they are strictly static in nature. Here we describe a novel variant of these classic illusions that incorporates stimulus dynamics and motion. We have found that adding various sources of motion to the stimulus can strongly modulate the magnitude of the contextual effects on perceived size. Thus a complete understanding of how and when contextual factors are integrated with primary-size cues needs to account for the effects of image dynamics. We propose a framework for understanding the effects of stimulus motion based on uncertainty in the neural representation of an object's image size.

THE DYNAMIC ILLUSORY SIZE-CONTRAST EFFECT

Figure II.27-1 shows a schematic of the illusion, which is viewable in its full form in Video II.27-1. In this illusion, the size of a black target bar appears to shrink as a surrounding box grows and to grow as the surrounding box shrinks (Fig. II.27-1, right). In fact, the target bar never changes its physical length. As the magnitude of the illusion primarily depends on the dynamic change in the relative size of the target bar and the surrounding background, we have termed the effect Dynamic Illusory Size Contrast (DISC; Mruczek, Blair, & Caplovitz, 2014).

So what distinguishes the DISC effect from a simple size-contrast effect played out over time? As stated, although the DISC effect is primarily a size-contrast effect, it is critically dependent on stimulus motion. There are a number of simple observations that reveal the underlying interaction between size contrast and stimulus motion. With respect to the necessity of relative size, when the surrounding box is unchanging so that there is no relative size change between the surrounding box and the target bar, there is no illusion (Video II.27-2). Thus motion dynamics induced by changes in target position and eye movements are, by themselves, ineffective at altering perceived size. Additionally, a relative size change between the surround box and the target bar without additional stimulus dynamics is equally ineffective at inducing a perceived size change (Video II.27-3).

In other words, the expanding box is necessary but not sufficient for generating an illusory percept. Rather, it is the interaction between the relative size change and other dynamic stimulus motion that drives the DISC effect. Movement of an object's image on the retina (e.g., dynamic stimulus motion) can arise from eye movements or movement of an object in real-world coordinates. Either or both of these sources of motion, when combined with the relative size change, are sufficient to induce the illusion. We have found that, taken in isolation from each other, stimulus motion is more potent than eye movements in modulating the perceived size of the target bar. It is the requirement of stimulus dynamics in addition to a relative size change that distinguishes the DISC effect from the static Delboeuf and Ebbinghaus illusions described earlier.

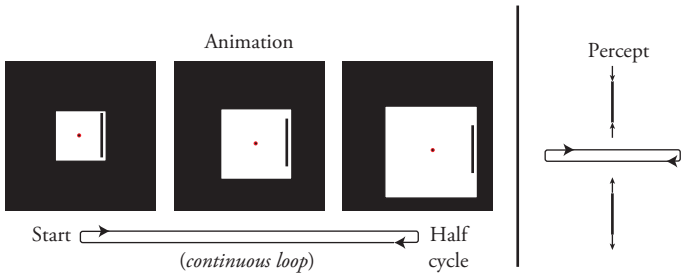
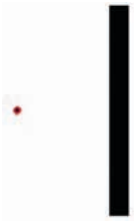


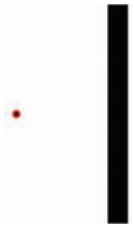
Figure II.27-1. A schematic of the DISC effect. The length of the black target bar appears to shrink as the surrounding box grows and vice versa (right panel). Although a corresponding illusion can be seen by comparing the length of the target bar in the two extreme frames, the dynamic version is subjectively more compelling (see Video II.27-1).



Video II.27-1. The DISC effect. The black target bar appears to shrink and expand as the surrounding box expands and shrinks. Movies are best viewed in looped mode.



Video II.27-2. If there is no relative size change between the surrounding box and the target bar, then the size of the target bar does not appear to change over time. Movies are best viewed in looped mode.



Video II.27-3. If the only dynamic change is the relative size change between the surrounding box and the target bar, then the size of the target bar does not appear to change over time, or it appears to change very little compared to the effect in Video II.27-1 Movies are best viewed in looped mode.

The perceived size of an object must be constructed by integrating information from the retinal image with other sources of information such as physical (or perceived) distance (Berryhill, Fendrich, & Olson, 2009; Boring, 1940; Emmert, 1881; Ponzo, 1911), and the relative size of different objects in a scene (Coren & Girgus, 1978; Delboeuf, 1865; Titchener, 1901), among factors. Changes in retinal image size, relative size, and perceived depth may lead to changes in perceived size, if such cues are strong enough. The DISC effect demonstrates that the contribution of relative size carries more weight under conditions in which both the target bar and contextual surround are physically moving or are perceived to be moving.

AN INFORMATION UNCERTAINTY HYPOTHESIS

Object movement is an inherent aspect of our natural environments. Thus dynamic illusions like the DISC effect can tell us something important about the conditions under which contextual influences contribute to size perception. We have proposed that the visual system weights the many different sources of information that contribute to size perception based on the level of uncertainty in the representation of that information (Mruczek et al., 2014). In the case of the DISC effect, stimulus motion leads the representation of image size to become more uncertain. This is because as the projected image moves across the retina or is tracked with imperfect pursuit eye movements, it will smear and deform over time. As such, the projected image will be less informative about physical size of the object it represents. According to the information uncertainty hypothesis, this will lead the contribution of image size to be reweighted downward, thereby increasing the relative contribution of other stimulus cues (i.e., relative size) to perceived size and thereby resulting in the DISC effect. Target motion and eye movements both serve to increase noise at the level of the retinal image, allowing relative size changes to alter perceived size of the target bar, as in the case in Video II.27-1. In contrast, when the retinal image is stable, as in Video II.27-3, relative size cues are essentially ignored. This relatively small effect likely reflects the fact that the weight for retinal image size is by default relatively high. Thus a representation of image size that is stable and unchanging over time is sufficient to counteract effects of the lower-weighted relative size.

Consistent with the information uncertainty hypothesis, we have found that a moving target and fixed-eyes version of the DISC stimulus produced a stronger illusion than a stationary target with moving eyes, even though the two conditions were matched for their retinal stimulation. Presumably, the visual system can partially account for relative motion between the target and eyes resulting from an eye movement using an efference copy (Bridgeman, 1995; Wurtz & Sommer, 2004) of the eye-movement command (leading to less uncertainty), whereas this is not possible for the same relative motion resulting from stimulus motion alone (leading to more uncertainty).

Some observations from the static versions of the Delboeuf and Ebbinghaus illusions are also consistent with the information uncertainty hypothesis. For example, decreasing the lightness contrast of the central circle magnifies the size-contrast effect in the large-surround versions of the Delboeuf and Ebbinghaus illusions, as well as the size-assimilation effect in the small-surround version of the Delboeuf illusion (Jaeger & Long, 2007; Jaeger & Pollack, 1977; Weintraub, Wilson, Greene, & Palmquist, 1969). Depicting the central circle in the Delboeuf illusion as a dashed line has the same result (Coren & Girgus, 1978; Oyama, 1962; Weintraub et al., 1969). It may be the case that, like stimulus motion, decreasing lightness contrast or using a dashed line for the target circle decreases the certainty of the representation of image size, lowers its corresponding contribution to perceive size, and thereby magnifies the contextual relative size effects. Finally, it almost goes without saying that image size alone is insufficient to accurately represent the physical size of a viewed object and that it must be integrated with a representation of physical (or perceived) viewing distance. Although we have investigated uncertainty only at the level of the image-size representation, the information uncertainty hypothesis makes the same predictions regarding the strength of contextual influences if there is increased uncertainty in the representation of viewing distance. Indeed, stronger illusory effects have been reported for the Ebbinghaus

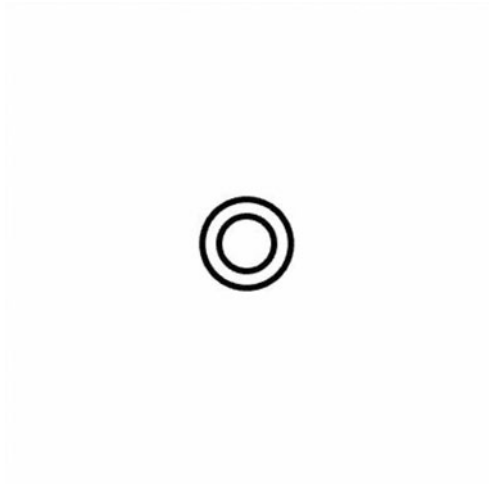
illusion (Song, Schwarzkopf, & Rees, 2011) and the Müller-Lyer illusion (Howard & Wagner, 1973) during monocular viewing conditions, which restrict several cues to physical viewing distance.

THE DYNAMIC DELBOEUF AND EBBINGHAUS ILLUSIONS

There is nothing inherently special about the stimulus configuration for the DISC effect shown in Figure II.27-1 and Video II.27-1. Indeed, the same effect is easily demonstrated using the Delboeuf (Videos II.27-4 and II.27-5) or Ebbinghaus (Fig. II.27-2, Videos II.27-6 and II.27-7; Mruczek et al., 2015) configurations.

These videos demonstrate the surprising result that a smooth transition in the size of the inducing surround has a negligible effect on the perceived size of a stationary target circle (Videos II.27-4 and II.27-6). In contrast, the addition of a translation in the entire stimulus revives the illusion (Videos II.27-5 and II.27-7). Moreover, the dynamic versions of these illusions are subjectively more compelling than the corresponding static versions (see Figures II.28-1 and II.28-2 in chapter 28 of this volume).

We have recently conducted a series of empirical studies examining the DISC effect using the Ebbinghaus configuration (Mruczek et al., 2015). The observations drawn from the experiments are consistent with the information uncertainty hypothesis outlined previously. For examples, the stationary (Video II.27-6) and moving (Video II.27-7) versions of the dynamic Ebbinghaus illusion are largely matched for their retinal stimulation when gaze is directed to the central target. However, in the later condition, the target circle is physically moving and must be tracked with eye movements. Both of these factors will lead to noise in the representation of image size of the target circle because imperfect smooth pursuit will lead to smearing and distortions of the projected image over time. This allows for a more influential effect of the surrounding inducers (Fig. II.27-2, right).



Video II.27-4. The "stationary" dynamic Delboeuf illusion. Perhaps surprisingly, a smooth transition in the size of the outer circle has little effect on the perceived size of the central circle. Compare with Video II.27-5. Movies are best viewed in looped mode.



Video II.27-5. The "moving" dynamic Delboeuf illusion. When the entire stimulus is translating in space, the central circle appears to change size over the course of the animation, shrinking as the outer circle grows and growing as the outer circle shrinks. Compare with Video II.27-4. Movies are best viewed in looped mode.

Under the information uncertainty hypothesis, other manipulations that increase uncertainty in the image-size representation are predicted to lead to a stronger illusion. For example, fixating the upper left inducer of Video II.27-6, thereby positioning the target circle in the lower acuity retinal periphery, revives the illusion. Finally, doing the same in Video II.27-7 leads to the largest subjective effect on the perceived size of the central target, which is consistent with the additive effects of eye movements, peripheral viewing, and target motion on the uncertainty in the target's image size. The complement of these effects can be observed in the original DISC stimulus (a target bar surrounded by a box; Video II.27-1), in which fixating the target bar rather than the fixation point greatly diminishes the illusion magnitude.

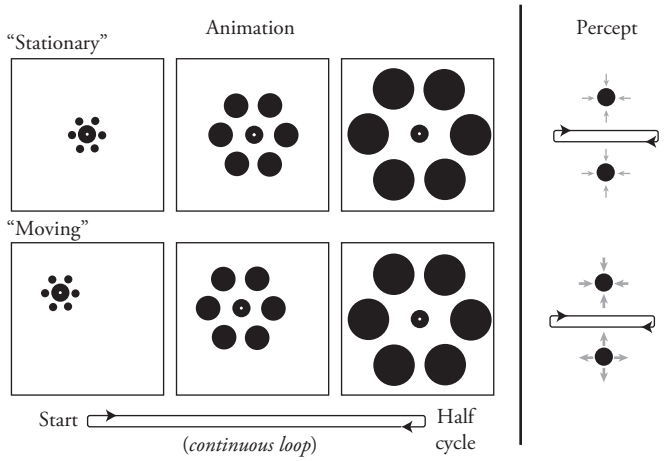


Figure II.27-2. A schematic of the dynamic Ebbinghaus illusion. The size of the central circle appears to shrink as the surrounding inducers grow and vice versa. When fixating the central circle (white dot), the illusory effect is much stronger (right) when the entire stimulus is moving across the screen (bottom, Video II.27-7) compared to when it is stationary (top, Video II.27-6).



Video II.27-6. The "stationary" dynamic Ebbinghaus illusion. Perhaps surprisingly, a smooth transition in the size of the inducers has little effect on the perceived size of the central circle. Compare with Video II.27-7. Movies are best viewed in looped mode.



Video II.27-7. The "moving" dynamic Ebbinghaus illusion. When the entire stimulus is translating in space, the central circle appears to change size over the course of the animation, shrinking as the inducers grow and growing as the inducers shrink. Compare with Video II.27-6. Movies are best viewed in looped mode.

CONCLUSIONS

Fundamental questions regarding how the brain constructs perceived size from multiple cues in the visual environment remain unanswered, and many potential contributing factors have not been fully explored. The DISC effect expands upon a century of work using the static Delboeuf and Ebbinghaus illusions by demonstrating how uncertainty in an object's image-size representation can increase the weight of relative size comparisons in the construction of its perceived size. In an environment that is inherently dynamic, these factors are likely to play an important role in everyday size judgments.

REFERENCES

Berryhill, M. E., Fendrich, R., & Olson, I. R. (2009). Impaired distance perception and size constancy following bilateral occipitoparietal damage. *Experimental Brain Research*, *194*(3), 381–393.

Boring, E. G. (1940). Size constancy and Emmert's law. *American Journal of Psychology*, *53*(2), 293–295.

Bridgeman, B. (1995). A review of the role of efference copy in sensory and oculomotor control systems. *Annals of Biomedical Engineering*, *23*(4), 409–422.

Burton, G. (2001). The tenacity of historical misinformation: Titchener did not invent the Titchener illusion. *History of Psychology*, *4*(3), 228–244.

Coren, S., & Girgus, J. S. (1978). *Seeing is deceiving: The psychology of visual illusions*. Hillsdale, NJ: Lawrence Erlbaum.

Delboeuf, M. J. (1865). Seconde note sur de nouvelles illusions d'optique: Essai D'une theorie psychophysique de la manière dont l'oeil apprécie les grandeurs [Second note on new optical illusions: Essay on a psychophysical theory concerning the way in which the eye evaluates sizes]. *Bulletins de l'Académie Royale des Sciences, Letters et Beaux-arts de Belgique*, *20*(2d Ser.), 70–97.

Emmert, E. (1881). Grössenverhältnisse der nachbilder. [Size relationships of afterimages]. *Klinische Monatsblätter für Augenheilkunde und für augenärztliche Fortbildung*, *19*, 443–450.

Howard, R. B., & Wagner, M. (1973). The role of contour and location mechanisms in the Müeller-Lyer illusion. *Bulletin of the Psychonomic Society*, *2*(4), 235–236.

Jaeger, T., & Long, S. (2007). Effects of contour proximity and lightness on Delboeuf illusions created by circumscribed letters. *Perceptual and Motor Skills*, *105*(1), 253–260.

Jaeger, T., & Pollack, R. H. (1977). Effect of contrast level and temporal order on the Ebbinghaus circles illusion. *Perception & Psychophysics*, *21*(1), 83–87.

Mruczek, R. E. B., Blair, C. D., & Caplovitz, G. P. (2014). Dynamic illusory size contrast: A relative-size illusion modulated by stimulus motion and eye movements. *Journal of Vision*, *14*(3). Advance online publication. doi: 10.1167/14.3.2

Mruczek, R. E. B., Blair, C. D., Strother, L., & Caplovitz, G. P. (2015). The dynamic Ebbinghaus: motion dynamics greatly enhance the classic contextual size illusion. *Front Hum Neurosci*, *9*, 77.

Nicolas, S. (1995). Joseph Delboeuf on visual illusions: A historical sketch. *The American Journal of Psychology*, *108*(4), 563–574.

Oyama, T. (1962). The effect of hue and brightness on the size-illusion of concentric circles. *The American Journal of Psychology*, *75*, 45–55.

Ponzo, M. (1911). Intorno ad alcune illusioni nel campo delle sensazioni tattili sull'illusione di Aristotele e fenomeni analoghi. [On some illusions in the field of tactile sensations on the illusion of Aristotle and similar phenomena]. *Archive für die Gesamte Psychologie*, *16*, 307–345.

Song, C., Schwarzkopf, D. S., & Rees, G. (2011). Interocular Induction of illusory size perception. *BMC Neuroscience*, *12*, 27. doi:10.1186/1471-2202-12-27

Thiéry, A. (1896). Über geometrisch-optische täuschungen [On geometric-optical illusions]. *Philosophische Studien*, *12*, 67–126.

Titchener, E. B. (1901). *Experimental psychology: A manual of laboratory practice*. New York: Macmillan.

Weintraub, D. J., Wilson, B. A., Greene, R. D., & Palmquist, M. J. (1969). Delboeuf illusion: Displacement versus diameter, arc deletions, and brightness contrast. *Journal of Experimental Psychology*, *80*(3), 505–511.

Wurtz, R. H., & Sommer, M. A. (2004). Identifying corollary discharges for movement in the primate brain. *Progress in Brain Research*, *144*, 47–60.

Chapter 28
Size Contrast and Assimilation in the Delboeuf and Ebbinghaus Illusions

Ryan E. B. Mruczek, Christopher D. Blair, Lars Strother, and Gideon P. Caplovitz

INTRODUCTION

Interacting with objects in our environment requires accurate neural representations of their physical size. Representation of an object's size requires the integration of on object's retinal image size with its physical (or perceived) distance (Berryhill, Fendrich, & Olson, 2009; Boring, 1940; Emmert, 1881; Ponzo, 1911). In addition, contextual factors, such as the relative size of different objects in a scene (Coren & Girgus, 1978; Delboeuf, 1865; Titchener, 1901), may also influence perceived object size. The Delboeuf illusion (Fig. II.28-21) and Ebbinghaus illusion (Fig. II.28-22) are two classic size illusions that demonstrate the contextual effects of surrounding visual information on the perceived size of an object. Under different conditions, these size illusions produce either *size contrast*, for which the size of a target object may be perceived to be more distinct from the surrounding objects than it really is, or *size assimilation*, for which the size of a target object may be perceived to be more similar to the surrounding objects than it really is.

The Delboeuf and Ebbinghaus illusions each have a rich history in the scientific study of perception and continue to contribute to our understanding of our ability to visually perceive the size of objects. In the following sections, we provide a brief history and overview of the primary observations for these classic illusions and review some theories regarding size contrast and size assimilation. Given the sheer number of studies investigating these illusions spanning more than a century, our goal is not to address all lines of research, theories, and applications but rather to give a representative sampling.

THE DELBOEUF ILLUSION

In the Delboeuf illusion (Fig. II.28-1), the perceived size of a target circle can be modulated by the presence of a larger concentric circle. If the outer circle is much larger than the target (Fig. II.28-1a), the target's size will be underestimated (i.e., size contrast). However, if the outer circle is only slightly larger than the target (Fig. II.28-1b), the target's size will be overestimated (i.e., size assimilation). Nicolas (1995) notes that Delboeuf first presented the illusion in 1865 (Delboeuf, 1865) in support of his overarching theory of visual illusions based on eye movement effort, one of the earliest general theories of geometric visual illusions. Under this account, for example, the eyes are drawn to the outer circle when it is close to the inner circle, and thus the eyes physically move over a greater distance when scanning the inner circle. The greater muscular effort of the eyes alters the perceived size of the central circle.

Since its introduction, the Delboeuf illusion has been the subject of numerous studies, and the phenomenology of both the assimilation and contrast effects have been well characterized. Assimilation effects are observed with closely spaced circles, with the maximum overestimation occurring when the ratio of the outer to inner circles is about 3:2 (Ikeda & Obanai, 1955; Morinaga, 1955; Piaget, Lambercier, Boesch, & von Albertini, 1942). Contrast effects are seen for widely space circles, with the maximum underestimation occurring when the same ratio is about 5:1 (Oyama, 1960). Although viewers are often directed to the apparent size change for the central circle, corresponding

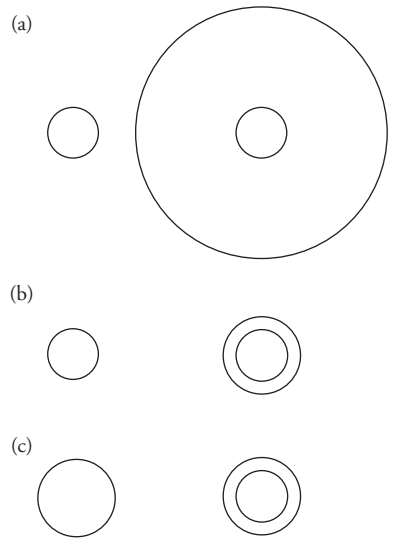


Figure II.28-1. The Delboeuf illusion. (a) The central circle appears smaller when surrounded by a concentric circle that is much larger, a size-contrast effect. (b) The central circle appears larger when surrounded by a concentric circle that is slightly larger, a size-assimilation effect. (c) The perceived size of the outer circle is also altered in these configurations.

assimilation or contrast effects are also seen for the outer circle (Oyama, 1960; Fig. II.28-1c). Additionally, although the most common form of the Delboeuf illusion consists of concentric circles, the illusion works for other matched shapes as well (Coren & Girgus, 1978; Robinson, 1972), and a variety of objects appear larger when they are closely encircled (Jaeger, 2007). The recently reported binding ring illusion (McCarthy, Kupitz, & Caplovitz, 2013) demonstrates assimilation effects on perceived size can also be observed when the target and contextual elements are superimposed.

In addition to the ratio of the sizes of the concentric circles, the magnitude of the Delboeuf illusion is also affected by the saliency of the constituent elements and the age of the viewer. Increasing the contrast of the outer circle or decreasing the contrast of the inner circle magnifies the effects and vice versa (Jaeger & Long, 2007; Oyama, 1962; Weintraub & Cooper, 1972; Weintraub, Wilson, Greene, & Palmquist, 1969). Depicting one of the circles as a dashed line results in the same modulatory effects as a decrease in contrast (Coren & Girgus, 1978; Oyama, 1962; Weintraub et al., 1969). Finally, the Delboeuf illusion decreases in magnitude with age (Santostefano, 1963; Wohlwill, 1960), and the main effect and its modulation by contrast are particularly potent for young children (Weintraub & Cooper, 1972).

THE EBBINGHAUS ILLUSION

In the Ebbinghaus illusion, the size of a central circle is misjudged depending on the size of a set of nonconcentric surrounding circles, or inducers (Fig. II.28-2). A circle surrounded by large inducers (Fig. II.28-2a) appears smaller than when the same target is surrounded by small inducers (Fig. II.28-2b). Burton (2001) notes that although this illusion is sometimes referred to as the Titchener circles illusions, a reference to a textbook that included a reprint of the illusion (Titchener, 1901) shows that Ebbinghaus deserves the credit as its originator (see also Boring, 1961). However, the fact that Ebbinghaus never explicitly published the illusion has led to confusion in the literature as to its origin and appropriate label. Regardless, like the Delboeuf illusion, the Ebbinghaus illusion has drawn the attention of many researchers.

Since its introduction, the Ebbinghaus illusion has commonly been interpreted as a prototypical size-contrast illusion in that the size of a central circle appears less like the surrounding circles than it really is. Consistent with a contrast mechanism, as the size of the inducers increases, the perceived size of the target circle decreases (Girgus, Coren, & Agdern, 1972; Massaro & Anderson, 1971; Roberts, Harris, & Yates, 2005). However, the contrast mechanism is not as straightforward as it might first seem. A reversal from overestimation to underestimation does not occur when the target and inducers are of equal size but rather when the target circle is about 77% as large as the inducers (Zigler, 1960). Thus it appears that the presence of any inducers may lead to an overestimation in target size, but this is counteracted by a size-contrast mechanism for larger inducers.

More directly, size contrast is probably not the only mechanism at play in the Ebbinghaus illusion. In many studies, inducer size is confounded with other stimulus parameters. The large-inducer condition often has fewer inducers with greater eccentricities and a larger overall extent of the stimulus, whereas the small-inducer condition often has more inducers that are closely spaced leading to a more complete surround. Research has demonstrated that each of these factors contributes to the illusory effects in the Ebbinghaus illusion. Increasing inducer eccentricity while holding inducer size constant (Fig. II.28-2c) causes the central circle to appear smaller (Ehrenstein & Hamada, 1995; Girgus et al., 1972; Jaeger & Grasso, 1993; Massaro & Anderson, 1971; Morinaga & Noguchi, 1966). Increasing the number of inducers, and thus the completeness of the surrounding context, leads to a stronger effect (Massaro & Anderson, 1971; Morinaga, 1956). However, even when these factors are held constant, changing the size of the inducers still leads to a size-contrast effect (Roberts et al., 2005). Roberts et al. argue that the two primary factors that determine the strength of the Ebbinghaus illusion are the size of the inducers and the separation between the target and the inducers. All together, it is likely that multiple mechanisms are involved and that they sometimes work together and sometimes in opposition (Girgus et al., 1972; Weintraub, 1979; Weintraub & Schneck, 1986).

In one sense, the Ebbinghaus illusion is more complex than the Delboeuf illusion in that there are a multitude of stimulus factors that affect the illusion magnitude. However, multiple researchers have pointed out similarities between the two illusions (Cooper & Weintraub, 1970; Girgus et al., 1972). In particular, inducer size, inducer eccentricity, and inducer spacing all affect the extent and completeness of the surrounding context. If we consider the extent of the Ebbinghaus stimulus as similar to the outer circle in the Delboeuf illusion, then we could interpret the illusory percepts as size contrast or size assimilation accordingly.

Similarity between the objects that comprise the central target and the surrounding context may also modulate the Ebbinghaus illusion. For example, a square surrounded by large squares appears smaller than the same square surrounded by large circles (Vuk & Podlesek, 2005). In general, a weaker illusion occurs when the target and inducers are less similar in terms of their physical shape (Coren & Miller, 1974) or perimeters (Choplin & Medin, 1999) and perhaps even their semantic relatedness (Coren & Enns,

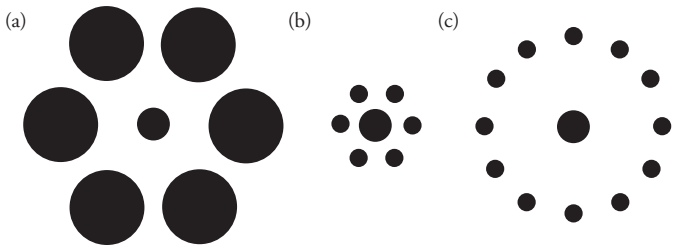


Figure II.28-2. The Ebbinghaus illusion. The center circle appears smaller when the inducers are large (a) compared to small (b). In addition to size contrast, inducer eccentricity also contributes to the perceived size of the central circle. The center circle appears larger for near inducers (b) compared to far inducers (c).

1993). However, consistent similarity effects have not been observed for all shapes (Rose & Bressan, 2002).

Like the Delboeuf illusion, the magnitude of the Ebbinghaus illusion is influenced by saliency and the age of the viewer. The effects of lightness contrast, however, are more complicated than for the Delboeuf illusion. Similar to the Delboeuf illusion, increasing the contrast of the surrounding circles or decreasing the contrast of the inner circle magnifies the overestimation effect in the presence of small inducers (Jaeger & Grasso, 1993; Jaeger & Pollack, 1977). On the other hand, the same manipulation decreases the underestimation effect in the presence of large inducers (Jaeger & Grasso, 1993; Jaeger & Pollack, 1977). Thus increasing the contrast of the inducers relative to the target may always have the effect of increasing the perceived size of the target. Age-related changes in illusion magnitude also differ between the Delboeuf and Ebbinghaus illusions. Sensitivity to size contrast in the Ebbinghaus illusion gradually develops in children, at least until the age of 10 (Doherty, Campbell, Tsuji, & Phillips, 2010; Weintraub, 1979), as do the semantic-related effects (Zanuttini, 1996).

MODELS AND THEORIES OF SIZE CONTRAST AND ASSIMILATION

As reviewed here, evidence suggests both similarities and differences between the Delboeuf and Ebbinghaus illusions. Descriptive models incorporating both illusions (e.g., Weintraub & Schneck, 1986) have emphasized the existence of multiple processes, including contour attraction (Jaeger, 1978; Jaeger & Grasso, 1993; Jaeger & Pollack, 1977; Rose & Bressan, 2002; Weintraub, 1979; Weintraub & Schneck, 1986), relative size contrast (Girgus et al., 1972; Massaro & Anderson, 1971; Roberts et al., 2005), stimulus extent and framing effects (Girgus et al., 1972; Roberts et al., 2005), and similarity (Choplin & Medin, 1999; Coren & Enns, 1993; Coren & Miller, 1974). But exactly how and why are these relative comparisons of size contrast and size assimilation instantiated?

First, it is worth reiterating some observations suggesting that the relative size comparison is instantiated at a relatively late stage of perceptual processing and thus cannot be explained by factors such as optic blur or lateral interactions at the level of the retina (Coren, 1999; Coren & Girgus, 1978). Although these mechanisms may contribute to size contrast under specific conditions, they are at odds with size assimilation (Coren & Girgus, 1978). Additionally, the fact that the Ebbinghaus illusion can be driven by a stimulus array in which the target and inducers are all the same physical size, but the apparent size of the inducers is manipulated by perceived depth via stereogram displays, indicates that the relative size comparison takes place after size constancy has been accounted for (Coren, 1971). The relative size comparison in the Ebbinghaus illusion also appears to take place after comparisons of similarity between the target and the surround (Choplin & Medin, 1999; Coren & Enns, 1993; Coren & Miller, 1974). An attentional selection process may also weight inputs to the relative size comparison. If both large and small inducers surround a target, the direction of the resulting illusion is predicted by which inducers are the focus of covert attention in a dual-task paradigm (Shulman, 1992). Finally, recent neuroanatomical and behavioral studies have linked the neural processing of size context to primary visual cortex (Fang, Boyaci, Kersten, & Murray, 2008; Murray, Boyaci, & Kersten, 2006; Schwarzkopf & Rees, 2013; Schwarzkopf, Song, & Rees, 2011; Song et al., 2013; Song, Schwarzkopf, & Rees, 2011; Sperandio, Chouinard, & Goodale, 2012). However, the representations of perceived size observed in V1 may arise due to feedback from even higher levels of visual cortex (Murray et al., 2006).

Although the illusory effects of the Delboeuf and Ebbinghaus illusions (as well as other related illusions) may arise from multiple and possibly distinct mechanisms, some theories have attempted to address both contrast and assimilation effects. Coren and Girgus (1978; Girgus & Coren, 1982) proposed a mechanistic theory to account for both size contrast and size assimilation effects in visual illusions. Their proposal builds on adaptation-level theory (Helson, 1964; Massaro & Anderson, 1971; Restle & Merryman, 1968; Wilson & Pressey, 1988), which states that object size is perceived in relation to a baseline level, which itself is determined from a weighted average of the size of all objects in a scene (or previously experienced). If the size of an object is close to the baseline, then assimilation occurs; if it is far from the baseline, then contrast occurs.

To formalize their theory, Coren and Girgus (1978; Girgus & Coren, 1982) proposed a pool and store model consisting of two underlying mechanisms that could account for these effects. First, an averaging or assimilating process works on all stimuli that are available in a single fixation. For example, in the near-inducer version of the Delboeuf, both circles are processed simultaneously so this assimilation process alters their perceived sizes. Thus pooling of stimulus properties takes place for incoming sensory information. A second process entailing stimulus contrast is needed when information about a stimulus must be gathered across multiple fixations. Under these conditions, information from one fixation must be stored in memory and compared with information gathered in subsequent fixations. To maintain the distinction between these two sources of information, the pool and store model proposes that a contrast mechanism emphasizes differences in the stored representation and the immediate sensory input. For example, in the far-inducer version of the Delboeuf, the target and contextual circles are processed sequentially so this contrast process alters their perceived sizes. The pool and store model also explains the primary observations from the Ebbinghaus illusion, assuming that the extent of the inducers provides a Delboeuf-like illusory effect (Cooper & Weintraub, 1970; Girgus et al., 1972).

In addition, the pool and store model makes predictions about when size-assimilation or size-contrast effects will be observed. Specifically, stimuli that are processed in a single glance will be subject to assimilation effects, whereas stimuli processed sequentially will be subject to contrast effects. Coren and Girgus (1978; Girgus & Coren, 1982) emphasize that both spatial and temporal proximity can determine which of the two processes dominates. From a temporal standpoint, their model can explain the observation that the size assimilation in the near-inducer

Delboeuf configuration switches to a size-contrast effect when the target and inducer are presented with a temporal offset (Cooper & Weintraub, 1970). However, Jaeger and Pollack (1977) report that the underestimation in the Ebbinghaus configuration, presumably a result of size contrast with large inducers, does not persist with temporally sequential presentations. They argue that their observations, instead, support a contour attraction model (Jaeger & Grasso, 1993; Jaeger & Pollack, 1977; Weintraub, 1979; Weintraub & Schneck, 1986).

Coren and Girgus (1978; Girgus & Coren, 1982) stressed that spatially distant regions of the visual field must be sampled via multiple fixations, but covert shifts of attention can occur in the absence of overt eye movements (Posner, 1980). Even in the absence of directed attention, individuals can process more than one item in a visual display (Hughes, Caplovitz, Loucks, & Fendrich, 2012; Mance, Becker, & Liu, 2012; Sperling, 1960). In addition, recent evidence supports the idea that the global impression or "gist" of a wide area of the visual field is extracted separately from the representation of any individual item (Albrecht & Scholl, 2010; Alvarez, 2011; Ariely, 2001; Kimchi, 1992; Navon, 1977; Oliva & Torralba, 2006). Unlike focal processing of individual objects, this global representation can carry the average properties of many stimuli, such as the mean size of an array of circles (Ariely, 2001; Chong & Treisman, 2003). This suggests that spatial proximity may also influence whether individual stimuli are incorporated into a pooled representation and thus subject to an averaging or assimilation process. However, it has been recently shown that a global impression of average size is extracted following the contextual influences observed in the Ebbinghaus illusion (Im & Chong, 2009). This would suggest that such ensemble encoding is not the mechanism by which the proposed pooling occurs. Even so, the pool and store model need not be dependent on multiple fixations per se.

Howe and Purves (2004) offer a very different explanation for the Ebbinghaus and Delboeuf illusions, one that is based on the statistical properties of our natural environment. They note that because the physical size of an object cannot be unambiguously determined from the retinal image size, size percepts should be constrained by the probability distribution reflecting the possible real-world sizes of objects that we encounter. If different-sized targets (e.g., a circle) are more or less likely to occur under different contexts (e.g., in the presence of concentric or nonconcentric surrounding objects), then this should influence our percepts in these contexts. Howe and Purves offer evidence in support of their hypothesis by measuring the statistical properties of real-world scene sections matching the Ebbinghaus and Delboeuf illusion configurations.

Overall, this theory provides an explanation of why size contrast and size assimilations like the Ebbinghaus and Delboeuf illusions occur and a mechanism by which this is neurally instantiated. As a result of statistical learning, our visual systems are biased toward interpreting certain retinal images in ways that are consistent with our previous experience. However, the theory cannot fully account for all of the observations reported here. For example, although the statistical learning hypothesis is consistent with the age-related increases in the magnitude of the Ebbinghaus illusion, presumably because younger children have not had sufficient time to acquire the experience necessary to tune their visual systems to the statistics of the natural world, the same logic is inconsistent with the age-related decreases in the magnitude of the Delboeuf illusion.

BEYOND SIZE PERCEPTION

The literature reviewed here largely focuses on characterizing the Delboeuf and Ebbinghaus illusions in order to inform our understanding of size perception. However, these illusions have also been utilized and applied to a variety of other research fields. For example, an individual's sensitivity to size contrast and size assimilation may be the result of different cognitive processing styles and perceptual strategies (Piaget, 1969; Witkin, 1967). As such, researchers have used the Ebbinghaus illusion in particular to explore differences across age groups (Doherty et al., 2010; Wapner & Werner, 1957), sexes (Phillips, Chapman, & Berry, 2004), cultures (de Fockert, Davidoff, Fagot, Parron, & Goldstein, 2007; Doherty, Tsuji, & Phillips, 2008), and cognitive impairments associated with psychological disorders such as autism (Happe, 1996; Ropar & Mitchell, 1999) and schizophrenia (Tibber et al., 2013; Uhlhaas, Silverstein, Phillips, & Lovell, 2004).

These illusions have also been used in an applied context. For example, the perceptual effects of the Delboeuf illusion have implications for eating behavior (McClain et al., 2013; Parrish & Beran, 2014; van Ittersum & Wansink, 2012; Wansink, van Ittersum, & Painter, 2006). People perceive a portion to be smaller when served on a larger plate, and this affects both self-portioned serving sizes and consumption. As such, Wansink and van Ittersum (2013) have proposed that eliminating large dishware would have beneficial effects for both consumers and businesses.

Mel Goodale and colleagues have utilized the Ebbinghaus illusion, among others, in a series of studies to investigate their perception-action model of primate vision (Goodale & Milner, 1992; Milner & Goodale, 2006). For example, they have shown that grip aperture, which scales with the size of a to-be-grasped target (Jeannerod, 1981), is not affected by the Ebbinghaus illusion (Aglioti, DeSouza, & Goodale, 1995; Haffenden & Goodale, 1998). Two circles that are perceptually different, because one is surrounded by large circles and the other by small circles, elicit the same grasping dynamics. In contrast, if the physical sizes of the central circles in the large-inducer and small-inducer stimuli are adjusted such that they are perceptually similar, grip aperture will reflect their physical difference. Such observations are consistent with a dichotomy of labor between a ventral vision-for-perception system, which is affected by the illusion, and a dorsal vision-for-action system, which provides real-time information for guiding actions that resists the illusion. Although some aspects of this work are controversial (for reviews, see Franz & Gegenfurtner, 2008; Goodale, 2011), it is an elegant example of the use of the Ebbinghaus illusion as a research tool to advance fundamental theories of visual organization.

CONCLUSIONS

Despite over a century of progress in visual psychology and neuroscience, many of the challenges in explaining assimilation and contrast effects on size perception remain highly relevant. Although unlikely to be due to the muscular effort of the eyes as posited by Delboeuf almost 150 years ago, the precise mechanisms that underlie size contrast and size assimilation remain unknown to this day. New discoveries about these and related effects are still being uncovered (Chapter 27). As such, classic size illusions such as the Delboeuf and Ebbinghaus continue to contribute to our understanding of the behavioral and neural basis of size perception, as well as vision more generally.

REFERENCES

Aglioti, S., DeSouza, J. F., & Goodale, M. A. (1995). Size-contrast illusions deceive the eye but not the hand. *Current Biology*, *5*(6), 679–685.

Albrecht, A. R., & Scholl, B. J. (2010). Perceptually averaging in a continuous visual world: Extracting statistical summary representations over time. *Psychological Science*, *21*(4), 560–567.

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Science*, *15*(3), 122–131.

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science*, *12*(2), 157–162.

Berryhill, M. E., Fendrich, R., & Olson, I. R. (2009). Impaired distance perception and size constancy following bilateral occipitoparietal damage. *Experimental Brain Research*, *194*(3), 381–393.

Boring, E. G. (1940). Size constancy and Emmert's law. *The American Journal of Psychology*, *53*(2), 293–295.

Boring, E. G. (1961). Letter to the editor. *Scientific American*, *204*(6), 18–20.

Burton, G. (2001). The tenacity of historical misinformation: Titchener did not invent the Titchener illusion. *History of Psychology*, *4*(3), 228–244.

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, *43*(4), 393–404.

Choplin, J. M., & Medin, D. L. (1999). Similarity of the perimeters in the Ebbinghaus illusion. *Perception & Psychophysics*, *61*(1), 3–12.

Cooper, L. A., & Weintraub, D. J. (1970). Delboeuf-type circle illusions: Interactions among luminance, temporal characteristics, and inducing-figure variations. *Journal of Experimental Psychology*, *85*(1), 75–82.

Coren, S. (1971). A size contrast illusion without physical size differences. *The American Journal of Psychology*, *84*(4), 565–566.

Coren, S. (1999). Relative contribution of lateral inhibition to the Delboeuf and Wundt-Hering illusions. *Perceptual and Motor Skills*, *88*(3 Pt 1), 771–784.

Coren, S., & Enns, J. T. (1993). Size contrast as a function of conceptual similarity between test and inducers. *Perception & Psychophysics*, *54*(5), 579–588.

Coren, S., & Girgus, J. S. (1978). *Seeing is deceiving: The psychology of visual illusions*. Hillsdale, NJ: Lawrence Erlbaum.

Coren, S., & Miller, J. S. (1974). Size contrast as a function of figural similarity. *Perception & Psychophysics*, *16*(2), 355–357.

de Fockert, J., Davidoff, J., Fagot, J., Parron, C., & Goldstein, J. (2007). More accurate size contrast judgments in the Ebbinghaus illusion by a remote culture. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(3), 738–742.

Delboeuf, M. J. (1865). Seconde note sur de nouvelles illusions d'optique: Essai d'une theorie psychophysique de la manière dont l'oeil apprécie les grandeurs [Second note on new optical illusions: Essay on a psychophysical theory concerning the way in which the eye evaluates sizes]. *Bulletins de l'Académie Royale des Sciences, Letters et Beaux-arts de Belgique*, *20*(2d serie), 70–97.

Doherty, M. J., Campbell, N. M., Tsuji, H., & Phillips, W. A. (2010). The Ebbinghaus illusion deceives adults but not young children. *Developmental Science*, *13*(5), 714–721.

Doherty, M. J., Tsuji, H., & Phillips, W. A. (2008). The context sensitivity of visual size perception varies across cultures. *Perception*, *37*(9), 1426–1433.

Ehrenstein, W. H., & Hamada, J. (1995). Structural factors of size contrast in the Ebbinghaus illusion. *Japanese Psychological Research*, *37*, 158–169.

Emmert, E. (1881). Grössenverhältnisse der nachbilder. [Size relationships of afterimages]. *Klinische Monatsblätter für Augenheilkunde und für augenärztliche Fortbildung*, *19*, 443–450.

Fang, F., Boyaci, H., Kersten, D., & Murray, S. O. (2008). Attention-dependent representation of a size illusion in human V1. *Current Biology*, *18*(21), 1707–1712.

Franz, V. H., & Gegenfurtner, K. R. (2008). Grasping visual illusions: Consistent data and no dissociation. *Cognitive Neuropsychology*, *25*(7), 920–950.

Girgus, J. S., & Coren, S. (1982). Assimilation and contrast illusions: Differences in plasticity. *Perception & Psychophysics*, *32*(6), 555–561.

Girgus, J. S., Coren, S., & Agdern, M. (1972). The interrelationship between the Ebbinghaus and Delboeuf illusions. *Journal of Experimental Psychology*, *95*(2), 453–455.

Goodale, M. A. (2011). Transforming vision into action. *Vision Research*, *51*(13), 1567–1587.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience*, *15*(1), 20–25.

Haffenden, A. M., & Goodale, M. A. (1998). The effect of pictorial illusion on prehension and perception. *Journal of Cognitive Neuroscience*, *10*(1), 122–136.

Happe, F. G. (1996). Studying weak central coherence at low levels: Children with autism do not succumb to visual illusions. A research note. *Journal of Child Psychology and Psychiatry*, *37*(7), 873–877.

Helson, H. (1964). *Adaptation-level theory: An experimental and systematic approach to behavior*. New York: Harper.

Howe, C. Q., & Purves, D. (2004). Size contrast and assimilation explained by the statistics of natural scene geometry. *Journal of Cognitive Neuroscience*, *16*(1), 90–102.

Hughes, H. C., Caplovitz, G. P., Loucks, R. A., & Fendrich, R. (2012). Attentive and pre-attentive processes in change detection and identification. *PLoS One*, *7*(8), e42851.

Ikeda, H., & Obanai, T. (1955). Figural after-effect, retroactive-effect and simultanious illusion. *Japanese Journal of Psychology*, *26*(4), 235–246.

Im, H. Y., & Chong, S. C. (2009). Computation of mean size is based on perceived size. *Attention, Perception & Psychophysics*, *71*(2), 375–384.

Jaeger, T. (1978). Ebbinghaus illusions: Size contrast or contour interaction phenomena? *Perception & Psychophysics*, *24*(4), 337–342.

Jaeger, T. (2007). Circumscribed shapes are enlarged: Is this a variation of the Delboeuf illusion? *Perceptual and Motor Skills*, *104*(3 Pt 2), 1116–1118.

Jaeger, T., & Grasso, K. (1993). Contour lightness and separation effects in the Ebbinghaus illusion. *Perceptual and Motor Skills*, *76*(1), 255–258.

Jaeger, T., & Long, S. (2007). Effects of contour proximity and lightness on Delboeuf illusions created by circumscribed letters. *Perceptual and Motor Skills*, *105*(1), 253–260.

Jaeger, T., & Pollack, R. H. (1977). Effect of contrast level and temporal order on the Ebbinghaus circles illusion. *Perception & Psychophysics*, *21*(1), 83–87.

Jeannerod, M. (1981). Intersegmental coordination during reaching at natural visual objects. In J. Long & A. Baddeley (Eds.), *Attention and performance* (pp. 153–168). Hillsdale, NJ: Lawrence Erlbaum.

Kimchi, R. (1992). Primacy of wholistic processing and global/local paradigm: A critical review. *Psychological Bulletin*, *112*(1), 24–38.

Mance, I., Becker, M. W., & Liu, T. (2012). Parallel consolidation of simple features into visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(2), 429–438.

Massaro, D. W., & Anderson, N. H. (1971). Judgmental model of the Ebbinghaus illusion. *Journal of Experimental Psychology*, *89*(1), 147–151.

McCarthy, J. D., Kupitz, C., & Caplovitz, G. P. (2013). The binding ring illusion: Assimilation affects the perceived size of a circular array [V2; Ref Status: Indexed, Http://F1000r.Es/12q]. *F1000Research*, *2*(58), 1–15.

McClain, A. D., van den Bos, W., Matheson, D., Desai, M., McClure, S. M., & Robinson, T. N. (2013). Visual illusions and plate design: The effects of plate rim widths and rim coloring on perceived food portion size. *International Journal of Obesity*, *38*(5), 657–662.

Milner, A. D., & Goodale, M. A. (2006). *The visual brain in action* (2d ed.). Oxford: Oxford University Press.

Morinaga, S. (1955). Optical ilusions and figural after-effects. Paper presented at the 19th meeting of the Japanese Psychological Association.

Morinaga, S. (1956). An examination of the conditions determining size-contrast. Paper presented at the 20th meeting of the Japanese Psychological Association.

Morinaga, S., & Noguchi, K. (1966). An attempt to unify the size-assimilation and size-contrast illusions. *Psychologische Forschung*, *29*(2), 161–168.

Murray, S. O., Boyaci, H., & Kersten, D. (2006). The representation of perceived angular size in human primary visual cortex. *Nature Neuroscience*, *9*(3), 429–434.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*(3), 353–383.

Nicolas, S. (1995). Joseph Delboeuf on visual illusions: A historical sketch. *The American Journal of Psychology*, *108*(4), 563–574.

Oliva, A., & Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research*, *155*, 23–36.

Oyama, T. (1960). Japanese studies on the so-called geometrical-optical illusions. *Psychologia*, *3*, 7–20.

Oyama, T. (1962). The effect of hue and brightness on the size-illusion of concentric circles. *The American Journal of Psychology*, *75*, 45–55.

Parrish, A. E., & Beran, M. J. (2014). When less is more: Like humans, chimpanzees (Pan Troglodytes) misperceive food amounts based on plate size. *Animal Cognition*, *17*(2), 427–434.

Phillips, W. A., Chapman, K. L., & Berry, P. D. (2004). Size perception is less context-sensitive in males. *Perception*, *33*(1), 79–86.

Piaget, J. (1969). *The mechanisms of perception* (G. N. Seagrine, Trans.). New York: Basic Books.

Piaget, J., Lambercier, M., Boesch, E., & von Albertini, B. (1942). Recherche sur le developpement des perceptions: I. Introduction a l'etude des perceptions chez l'enfant et analyse d'une illusion relative a la perception visuelle de cercles concentrique (Delboeuf). [Research on the development of perceptions: I. Introduction to the study of perception in children and analysis of a relative illusion in visual perception of concentric circles (Delboeuf)]. *Archives de Psychologie (Geneva)*, *43*(29), 1–107.

Ponzo, M. (1911). Intorno ad alcune illusioni nel campo delle sensazioni tattili sull'illusione di Aristotele e fenomeni analoghi. [On some illusions in the field of tactile sensations on the illusion of Aristotle and similar phenomena]. *Archive für die Gesamte Psychologie*, *16*, 307–345.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25.

Restle, F., & Merryman, C. T. (1968). An adaptation-level theory account of a relative-size illusion. *Psychonomic Science*, *12*(6), 229–230.

Roberts, B., Harris, M. G., & Yates, T. A. (2005). The roles of inducer size and distance in the Ebbinghaus illusion (Titchener circles). *Perception*, *34*(7), 847–856.

Robinson, J. O. (1972). *The psychology of visual illusions*. London: Hutchinson Education.

Ropar, D., & Mitchell, P. (1999). Are individuals with autism and Asperger's syndrome susceptible to visual illusions?. *Journal of Child Psychology and Psychiatry*, *40*(8), 1283–1293.

Rose, D., & Bressan, P. (2002). Going round in circles: Shape effects in the Ebbinghaus illusion. *Spatial Vision*, *15*(2), 191–203.

Santostefano, S. (1963). A developmental study of the Delboeuf illusion. *Perceptual and Motor Skills*, *17*(1), 23–29.

Schwarzkopf, D. S., & Rees, G. (2013). Subjective size perception depends on central visual cortical magnification in human V1. *PLoS One*, *8*(3), e60550.

Schwarzkopf, D. S., Song, C., & Rees, G. (2011). The surface area of human V1 predicts the subjective experience of object size. *Nature Neuroscience*, *14*(1), 28–30.

Shulman, G. L. (1992). Attentional modulation of size contrast. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *45*(4), 529–546.

Song, C., Schwarzkopf, D. S., Lutti, A., Li, B., Kanai, R., & Rees, G. (2013). Effective connectivity within human primary visual cortex predicts interindividual diversity in illusory perception. *Journal of Neuroscience*, *33*(48), 18781–18791.

Song, C., Schwarzkopf, D. S., & Rees, G. (2011). Interocular induction of illusory size perception. *BMC Neuroscience*, *12*, 27.

Sperandio, I., Chouinard, P. A., & Goodale, M. A. (2012). Retinotopic activity in V1 reflects the perceived and not the retinal size of an afterimage. *Nature Neuroscience*, *15*(4), 540–542.

Sperling, George. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, *74*(11), 1.

Tibber, M. S., Anderson, E. J., Bobin, T., Antonova, E., Seabright, A., Wright, B., … Dakin, S. C. (2013). Visual surround suppression in schizophrenia. *Frontiers in Psycholog,y 4*, 88. doi: 10.3389/fpsyg.2013.00088

Titchener, E. B. (1901). *Experimental psychology: A manual of laboratory practice*. New York: Macmillan.

Uhlhaas, P. J., Silverstein, S. M., Phillips, W. A., & Lovell, P. G. (2004). Evidence for impaired visual context processing in schizotypy with thought disorder. *Schizophrenia Research*, *68*(2–3), 249–260.

van Ittersum, K., & Wansink, B. (2012). Plate size and color suggestibility: The Delboeuf illusion's bias on serving and eating behavior. *Journal of Consumer Research*, *39*(2), 215–228.

Vuk, M., & Podlesek, A. (2005). The effect of inner elements of the context figures on the Ebbinghaus illusion. *Horizons of Psychology*, *14*(4), 9–22.

Wansink, B., & van Ittersum, K. (2013). Portion size me: Plate-size induced consumption norms and win-win solutions for reducing food intake and waste. *Journal of Experimental Psychology: Applied*, *19*(4), 320–332.

Wansink, B., van Ittersum, K., & Painter, J. E. (2006). Ice cream illusions: Bowls, spoons, and self-served portion sizes. *American Journal of Preventative Medicine*, *31*(3), 240–243.

Wapner, S., & Werner, H. (1957). *Perceptual development: An investigation within the framework of sensory-tonic field theory*. Worcester, MA: Clark University Press.

Weintraub, D. J. (1979). Ebbinghaus illusion: Context, contour, and age influence the judged size of a circle amidst circles. *Journal of Experimental Psychology: Human Perception and Performance*, *5*(2), 353.

Weintraub, D. J., & Cooper, L. A. (1972). Coming of age with the Delboeuf illusion: Brightness contrast, cognition, and perceptual development. *Developmental Psychology*, *6*(2), 187.

Weintraub, D. J., & Schneck, M. K. (1986). Fragments of Delboeuf and Ebbinghaus illusions: Contour/context explorations of misjudged circle size. *Perception & Psychophysics*, *40*(3), 147–158.

Weintraub, D. J., Wilson, B. A., Greene, R. D., & Palmquist, M. J. (1969). Delboeuf illusion: Displacement versus diameter, arc deletions, and brightness contrast. *Journal of Experimental Psychology*, *80*(3), 505–511.

Wilson, A. E., & Pressey, A. W. (1988). Contrast and assimilation in the Baldwin illusion. *Perceptual and Motor Skills*, *66*(1), 195–204.

Witkin, H. A. (1967). A cognitive-style approach to cross-cultural research. *International Journal of Psychology 2*(4), 233–250.

Wohlwill, J. F. (1960). Developmental studies of perception. *Psychological Bulletin*, *57*, 249–288.

Zanuttini, L. (1996). Figural and semantic factors in change in the Ebbinghaus illusion across four age groups of children. *Perceptual and Motor Skills*, *82*(1), 15–18.

Zigler, E. (1960). Size estimates of circles as a function of size of adjacent circles. *Perceptual and Motor Skills*, *11*(1), 47–53.

Chapter 29
The Occlusion, Configural Shape, and Shrinkage Illusions

Stephen E. Palmer and Karen B. Schloss

INTRODUCTION

Although the visual system is remarkably accurate in gauging the properties of most environmental objects under most ecologically reasonable circumstances, systematic visual illusions nevertheless occur in the perception of many visual properties. Among the best known and best understood of these are size illusions, most of which can be explained either by size contrast with nearby contextual objects (e.g., the Ebbinghaus illusion), by size assimilation to nearby contextual figures (e.g., the Delboeuf illusion), or by systematic errors in the perceived distance to the target objects (e.g., the moon illusion).

This chapter discusses an interesting suite of three related illusions of size and shape: the occlusion illusion, the configural shape illusion (CSI), and the shrinkage illusion. They are related to each other by virtue of the close commonality among the kinds of images that induce them (see Fig. II.29-1). Indeed, all three can be illustrated by displays that are perceived as a closer rectangular surface partly occluding a farther one. They are all the more fascinating because some of them work in opposite directions than others, making corresponding elements appear larger in some cases and smaller in others. A final factor of interest is that one of them—the occlusion illusion—concerns a perceptual illusion that appears to be based on quite a different mechanism from the classic size illusions mentioned earlier: namely, the possibility that the visual system "modally" (visibly) fills-in a thin strip of surface where it has been occluded by an opaque, closer surface.

To begin, consider Figures II.29-1a and II.29-1b, which show a larger gray and a smaller black rectangle, respectively. Figure II.29-1c shows the occlusion illusion using the same two rectangles but with the smaller black rectangle touching the gray one along a shared border. Notice that the black rectangle now appears noticeably larger than the physically identical one in Figure II.29-1b (Kanizsa, 1979; Kanizsa & Luccio, 1978; Micali, Giurissevich, & Serani, 1978; Palmer, Brooks & Lai, 2007; Palmer & Schloss, 2009). In fact, the occlusion illusion is a misperception of both shape and size, because the black target region appears selectively larger in one dimension (here its width), leading to differences in both the perceived size and shape of the black rectangle.

There is a second size/shape illusion present in the image shown in Figure II.29-1c but this time in perceiving the gray rectangle, which is seen as wider than the physically identical rectangle in Figure II.29-1a (Schloss, Fortenbaugh & Palmer, 2014). This illusion, which we call the configural shape illusion (CSI), is governed by different factors than the occlusion illusion, in part because it does not matter whether the gray rectangular target is occluded or seen as closer or farther than the black rectangular inducer. Rather, it is systematically influenced by factors that concern the strength of perceptual grouping between the target and inducer (stronger grouping generally produces a larger illusion) and the relative distance between the corresponding edges of the target and inducer.

The third illusion is interestingly related to both the occlusion illusion and the CSI. It occurs when a second smaller black rectangle is positioned on the opposite side of Figure II.29-1c so that the gray rectangle appears to occlude the central portion of a single centrally occluded black rectangle that is visible on both sides, as shown in Figure II.29-1d. Paradoxically, this image produces an illusion of the perceived size of the black rectangle that is *opposite* that of the one-sided occlusion illusion, making the centrally occluded black rectangle appear *smaller and narrower* than when it is fully visible (Fig. II.29-1e). It is therefore often called the shrinkage illusion (Kanizsa,

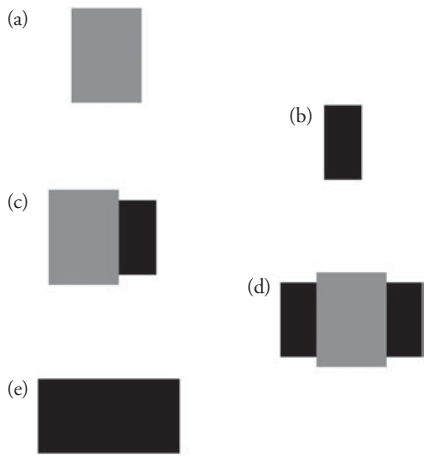


Figure II.29-1. Examples of the occlusion illusion (c vs. b), the CSI (c vs. a), and the shrinkage illusion (d vs. e). (Image by the authors for this chapter.)

1979). Its paradoxical nature stems from the fact that it seems to consist of two occlusion illusions (of the visible left and right portions of the black rectangle) plus a double-sided CSI (in which the gray rectangle appears wider than in Fig. II.29-1a). As such, one would expect the size of the centrally occluded black rectangle in Figure II.29-1d to be seen as substantially *wider* than it is, yet it produces a substantial illusion in the *opposite* direction, contracting the partly occluded target rectangle along the relevant dimension. This surprising reversal requires explanation.

Next we consider the three illusions in more detail: the factors that influence them, the mechanisms that might produce them, and the relations among them.

THE OCCLUSION ILLUSION

We begin by considering whether any of the three classic theories of size illusions—size contrast, size assimilation, or misperceived depth—provide an adequate explanation of the occlusion illusion and whether perceived occlusion is indeed critical in its occurrence. Figures II.29-2a to II.29-2d show variants of the occlusion illusion display (using target semicircles rather than rectangles) that differ in terms of the overall size of the gray rectangle and whether it appears to occlude the target semicircle. *Size contrast* fails as an explanation because the adjacent gray occluding rectangle in Figure II.29-2a is actually *larger* than the target. Size contrast thus would predict that the target should be perceived as *smaller* when it is occluded by the larger rectangle than when it is fully visible, but instead the target is perceived as larger. By similar logic, *size assimilation* fails as an explanation because the illusion works in the same direction (i.e., target appearing larger; see Fig. II.29-2c vs. Fig. II.29-2b) when the occluding gray rectangle is actually smaller than the target. Therefore, if relative size influences the occlusion illusion at all, it seems to be a minor factor. We also note that the magnitude of the occlusion illusion is largely invariant over luminance reversals (Palmer et al., 2007).

Perceived partial occlusion is critical, however. When the same rectangle as in Figure II.29-2a is positioned next to the black target rectangle, so that it does *not* appear to occlude the black target (Fig. II.29-2d), the illusion disappears. Further evidence of the crucial role of perceived occlusion comes from demonstration that the magnitude of the illusion varies with the strength of the cues to occlusion (Palmer et al., 2007). As illustrated in Figures II.29-3a to II.29-3e, again using semicircles rather than rectangles as targets, the illusion is largest with a fully visible occluder (Fig. II.29-3a), somewhat smaller with a strong illusory-figure occluder (Fig. II.29-3b), smaller still with a weaker illusory occluder (Fig. II.29-3c), and virtually absent with a control condition in which the same elements as in Figure II.29-3C are rotated such that there is no perception of an illusory surface occluding the target (Fig. II.29-3d). Such effects have been interpreted by Scherzer and Ekroll (2015) as indicating that visual percepts are experienced as modal whenever they are based on sufficiently conclusive sensory evidence, and otherwise as amodal.

An alternative account of the occlusion illusion can be given in terms of misperceived distance. The fact that the target is partly occluded implies that it is at least somewhat

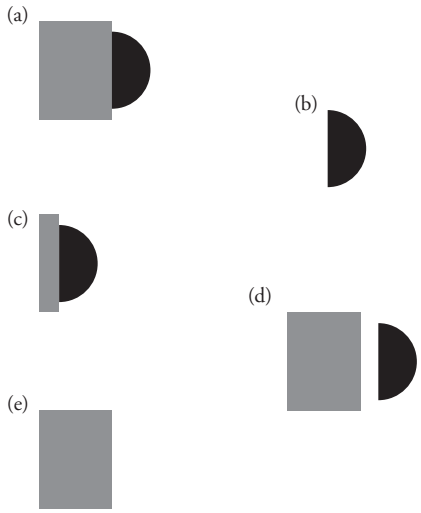


Figure II.29-2. Examples of the occlusion illusion that differ in rectangle size (a vs. c in comparison with b) and occlusion (a vs. d in comparison with b). Also shown are examples of the CSI that differ in occlusion (a vs. d in comparison with e). (Image by the authors for this chapter.)

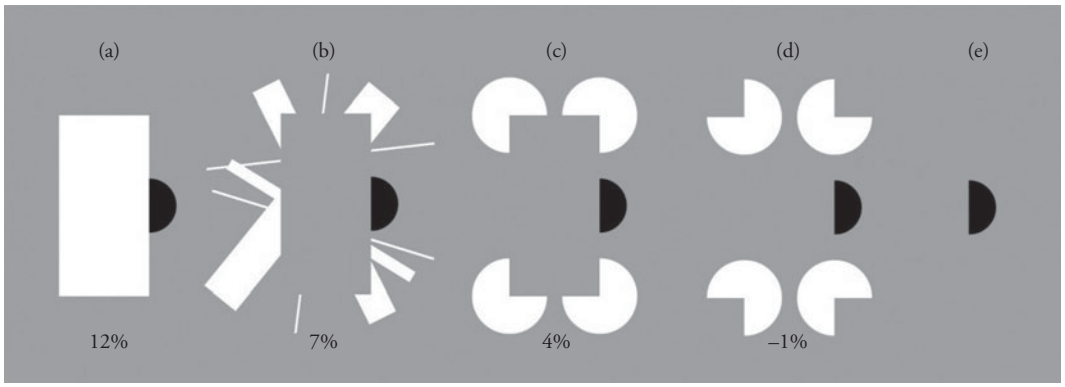


Figure II.29-3. Variations in magnitude for the occlusion illusion with different degrees of evidence for occlusion: a fully visible occluder (a vs. e), a densely specified illusory occluder (b vs. e), a sparsely specified illusory occluder (c vs. e), and a control condition with no illusory occluder (d) in comparison with a physically identical semicircular target (e). Numbers below each figure indicate the measured size of the illusion in percentages of horizontal extent. (Adapted from Palmer et al., 2007, Fig. 4.)

farther from the viewer, and, according to Emmert's law, this increase in perceived distance for images of the same retinal size, should make the target appear larger to an extent that depends on the additional perceived distance. This explanation is satisfying both because it is consistent with the crucial role of occlusion and because it appeals to a well-known principle underlying other size illusions, such as the moon illusion.

There is convincing evidence, however, that the apparent depth explanation is incorrect. Palmer and Schloss (2009) tested it by manipulating the relative binocular disparities of the target and occluding rectangle such that the target would be perceived at different distances behind and in front of the inducing rectangle. For each such display they measured both the perceived size and the perceived distance of the target. They found that increasing the stereoscopic distance to the target increased both its perceived distance and its perceived size, consistent with Emmert's law. They also found that occluding the target by the rectangle produced an additional increment in the target's perceived size—that is, an occlusion illusion—but with no evidence of a corresponding increase in its perceived distance. These results demonstrate that the occlusion illusion is not due to misperceived distance.

An additional implication of the apparent distance hypothesis is that the target should expand uniformly in size. Further evidence against the apparent distance hypothesis—and, indeed, any explanation that implies a uniform expansion in the target's size—comes from evidence that the occlusion illusion produces distortions in shape rather than in overall size (Palmer et al., 2007). The experimenters first found the points of subjective equality in the perceived area for the same partly occluded semicircular target (Fig. II.29-4a) with two series of fully visible standards. The "size-only" series differed solely in overall size (i.e., the different semicircles were uniform expansions/contractions of one another) whereas the "shape-change" series differed in both size and shape due to the addition (or subtraction) of a thin rectangular region along the occluded edge, as if slightly more (or less) of a semicircle-tipped "bullet-shaped" region were visible behind the occluder. After the points of subjective equality for both sequences were determined for each participant, they were presented with the occlusion illusion (Fig. II.29-3a) and the two standards that were matched with its target in size—say, the rightmost versions in Figures II.29-3b and II.29-3c—and were asked to choose the one that appeared more similar to the visible portion of the partly occluded target. They chose the shape-change version whose physical shape was different from that of the target reliably more often, supporting the conclusion that the occlusion illusion induces perceptual differences in both size and shape rather than just in size.

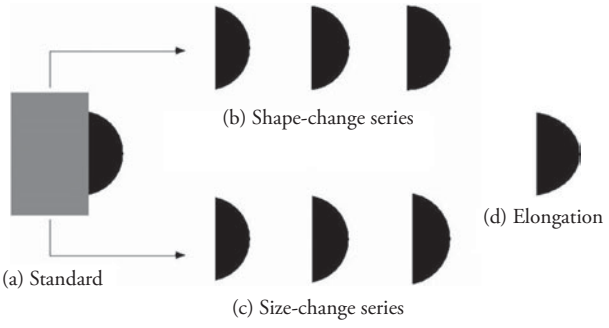


Figure II.29-4. Discriminating illusions of shape versus size in the occlusion illusion. The left version in each series is physically identical to the semicircular target in the standard illusion (a). The shape-changed version (b) that was matched for perceived area with the standard (in a) was judged to be more similar to the target (in a) than the size-changed version (c) that was similarly matched for perceived area. Shape elongation (d) does not appear to fit the nature of the shape change induced by the occluder. (Adapted from Palmer et al., 2007, Fig. 8.)

This conclusion is consistent with at least two theoretical interpretations. One is the partial modal completion (filling-in) hypothesis, described earlier, that the visual system effectively "fills in" a thin, visible strip along the occluded boundary of the target. Another is that the target is selectively elongated by global extension along the dimension toward the occluding rectangle. These two accounts differ in terms of the particular change in shape of the perceived target region: the selective elongation hypothesis implies that the curvature of the visible portion would change noticeably (see Fig. II.29-4d), whereas the partial modal completion hypothesis (Fig. II.29-4b) does not. Although we have not yet performed careful psychophysical measurements like those of Palmer et al. (2007), inspection strongly suggests that the curvature of the visible portion does not appear to be altered in the illusion as predicted by selective global extension. We therefore continue to believe that the account in terms of partial modal completion is the best explanation of the occlusion illusion currently available.

THE CONFIGURAL SHAPE ILLUSION

The CSI is another illusion evident in Figure II.29-1c. Here the shape of the target (now the gray rectangle) is systematically distorted by the presence of the attached black rectangle, which is distinct from, but perceptually grouped with, the target. The illusion is that the target rectangle is perceived as selectively elongated in a direction consistent with the configuration that includes the attached inducer, making the gray rectangle in Figure II.29-1c seem wider (and possibly shorter) than an identical rectangle viewed against a homogeneous background (Fig. II.29-1a).

The CSI seems, at first blush, to be closely related to the occlusion illusion, but, unlike its cousin, the CSI does not depend on occlusion. Indeed, when its magnitude is measured for inducers that are perceived as in front versus behind the target rectangle through two-dimensional occlusion cues, there is no difference. What does matter substantially, however, is how strongly the target and inducer are grouped together into a larger configuration. All of the factors known to increase grouping strength that have been tested for the CSI also increase the magnitude of the illusion, including good continuation, lightness similarity, proximity, and common fate (Schloss, Fortenbaugh, & Palmer, 2014).

To understand the mechanisms underlying the CSI, Schloss et al. (2014) measured CSI magnitude as a function

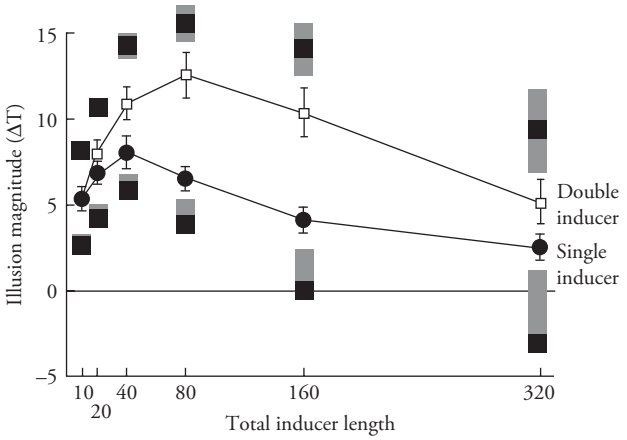


Figure II.29-5. CSI magnitude plotted as a function of total inducer length for one- and two-sided inducers. As the extent of the inducer (gray rectangles) increases relative to the target (black rectangles), the size of the illusory extension first increases dramatically, then decreases slowly. Two-sided inducers (double; top curve) produce systematically larger illusions than one-sided inducers (single; bottom curve) of the same total extent. (Adapted from Schloss et al., 2014, Fig. 6a.)

of variations in both the height of the inducer (the gray rectangles in Fig. II.29-5) and the height of the target rectangle (the black rectangles in Fig. II.29-5), with their widths being the same. As the inducer increased in height, the CSI first increased rapidly and then decreased slowly, as shown in the lower "single inducer" curve of Fig. II.29-5, but it never went below zero (i.e., an adjacent, external inducer always made the target appear larger). The maximum illusion occurred when a single inducer was about 20% to 30% of the target's size. This result suggests that the CSI is an illusion of assimilation in which the extent of the target is increased to some degree by the inclusion of the inducer in the same configuration, with greater grouping strength producing larger illusions.

One important theoretical issue is whether the mechanisms of the CSI operate globally in estimating the overall extent of the target or locally in estimating the positions of the target's edges while judging overall extent. To find out, Schloss et al. (2014) compared the magnitude of the one-sided CSI and a two-sided CSI in which the total extent of the one-sided inducer was the same as that of the pair of two-sided inducers. The results shown in Figure II.29-5 (upper curve) clearly show that the two-sided illusion is substantially larger than the one-sided illusion of equal physical extent for all inducer lengths tested. This finding implies that the CSI depends on errors in perceiving the positions of the edges of the target, which are displaced toward the corresponding edges of the whole configuration, rather than simply being an assimilation effect of the target on the total length of the configuration.

THE SHRINKAGE ILLUSION

The third and final illusion we discuss is often called the shrinkage illusion (e.g., Kanizsa, 1979). It is related to both of the illusions discussed previously but in a paradoxical way. In its canonical form, the shrinkage illusion occurs when a farther target shape is partly occluded in its center portion by a closer inducer such that the target is visible on both sides of the occluder (see Fig. II.29-6a). The illusion is that the amodally completed target behind the occluder is perceived to be noticeably *narrower* than a physically equivalent, unoccluded comparison figure (Fig. II.29-6c). Even though the inducer figure occludes the target in this form, occlusion itself plays only a minor role, if any, in the shrinkage illusion (Vezzani, 1999). One kind of evidence for this conclusion is that the shrinkage in a standard display (Fig. II.29-6a) is roughly the same as the shrinkage when the inducer is occluded by, rather than occluding, the target rectangle (Fig. II.29-6b).

The shrinkage illusion is particularly fascinating because it biases perception of the target's size and shape in the direction *opposite* to what would logically be predicted from considering illusions present in the perception of its three visible components. First, consider the left and right outer parts of the centrally occluded target in Figure II.29-6a. The occlusion illusion clearly implies that both portions should appear wider than they are: that is, the outer segments should be seen as horizontally expanded relative to physically equivalent segments that are not occluded. And, in fact, they are (see Fig. II.29-6a vs. Fig. II.29-6d). The fact that the shrinkage illusion display actually includes a double occlusion illusion clearly implies that the centrally occluded target should be perceived as substantially *wider* than the physically equivalent but unoccluded target (Fig. II.29-6c). Nevertheless, the entire target is perceived as smaller and narrower (Fig. II.29-6a vs. Fig. II.29-6c). The shrinkage of the target could still be explained, of course, if the central occluding rectangle were perceived to be substantially narrower than it actually is. In fact, however,

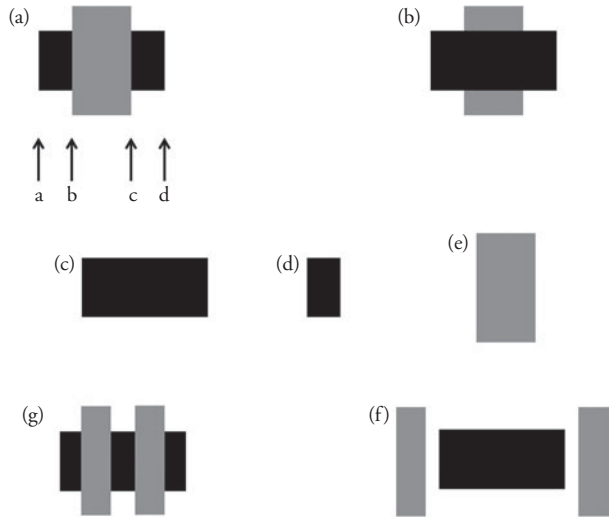


Figure II.29-6. The shrinkage illusion. In the shrinkage illusion, a center-occluded black rectangle (a) appears narrower than a physically identical rectangle that is fully visible (c) despite the fact that its two visible components (the left and right sides of a) appear wider than a physically identical version (d). The central occluder (in a) also appears wider than a physically identical version (e). The corresponding illusion appears in the opposite direction of expansion, however, if the inducing rectangle (in a) is split into two parts and moved outside of the target (f) but still produces shrinkage if the two inducing parts remain inside the target (g). (Image by the authors for this chapter.)

this occluding rectangle also appears wider than it is due to a double-sided CSI effect from the two flanking black segments (Fig. II.29-6a vs. Fig. II.29-6e). Thus all three visible component rectangles should be—and indeed are, at least individually—perceived to be wider than they actually are. How, then, could the centrally occluded black rectangle in Figure II.29-6a possibly be perceived as substantially narrower than it actually is (Fig. II.29-6c)?

The account we favor is surprisingly simple: people are just estimating different extents in the same display in the shrinkage illusion than they are when estimating its apparent components (i.e., in the portions of the display that constitute the occlusion illusion and CSI), and these different estimates are subject to different contextual influences. In judging the size of the occluding gray rectangle, for example, one must judge the distance from *b* to *c* in Figure II.29-6a, and the positions of these edges are subject to a *positive* (i.e., expansive) CSI due to the influence of the contextual edges at *a* and *d*, which are on the outside and thus tend to pull edges *b* and *c* farther apart visually. When judging the size of the partly occluded target rectangle, however, one must judge the distance from *a* to *d*, where the contextual edges at *b* and *c* are on the inside and thus produce a *negative* (i.e., contractive) CSI by effectively pulling edges *a* and *d* closer together visually. From this perspective, the combined extent of what would seem to be the three relevant visible regions comprising the illusory display—that is, the left visible part of the partly occluded target (*a* to *b*), the central occluding rectangle (*b* to *c*), and the right visible part of the partly occluded target (*c* to *d*)—is essentially irrelevant, because the required estimation of extent is made directly (*a* to *d*) and is influenced by different contextual elements and relations.

A variety of explanations for the shrinkage illusion have been proposed and evaluated experimentally. As reviewed by Verzanni (1999), they include various occlusion-based hypotheses involving mechanisms of amodal completion. These appear to be incompatible with the relative lack of difference between the illusions present in Figures II.29-6a and II.29-6b, however, because they must predict radical differences between Figures II.29-6a and II.29-6b. Other possibilities include size-contrast hypotheses (e.g., lateral inhibition of spatial frequency mechanisms; Zanforlin, 1981) and size-assimilation hypotheses (e.g., Luccio, 1981 following Pressey & Wilson, 1980).

It is worth emphasizing the close—though perceptually opposite—relations between the shrinkage illusion and the CSI. The shrinkage illusion makes the partly occluded target figure *contract*, because the edges of the inducer figure are *inside* those of the target figure. The CSI makes the partly occluded target figure *expand*, because the edges of the inducer figure are *outside* those of the target figure. These observations suggest that a unified account of both illusions may be possible in terms of an edge-based assimilation (Schloss et al., 2014). The difference in terms of the direction of the illusion (expansion in the CSI vs. contraction in the shrinkage illusion) thus may be governed by the placement of the supposedly irrelevant inducer edges relative to the relevant target edges.

An edge-based assimilation account for the shrinkage illusion becomes even more attractive when one considers the effects of other edge-based transformations in the basic shrinkage-illusion display. First, notice that if the partly occluding inducer is split into two parts that nevertheless remain inside the target rectangle (Fig. II.29-6f), the same contraction in the width of the rectangle is evident. Moving the inducers outside the target rectangle, however, does not merely eliminate the illusion but decisively reverses it (Fig. II.29-6g). Although edge-based assimilation may not be the only factor relevant to these effects, the fact that it clearly predicts the reversal is encouraging. Moreover, these displays make plain a further connection from the CSI and shrinkage illusions to the Müller-Lyer illusion and other related illusions of linear extent.

How might such edge-based assimilation effects arise from neural processing in the visual system? Edge-assimilation effects might come from spatial pooling across overlapping receptive fields that contain an irrelevant inducer edge as well as the relevant target edge. This type of edge-based spatial averaging model has previously been applied to account for illusions of extent for a line or a space flanked by contextual objects (e.g., Bulatov, Bertulis, Gutauskas, Mickiene, & Kadziene, 2010; Morgan, Hole, & Glennerster, 1990) and errors in the perceived location of dots (e.g., Prinzmetal, 2005). Such edge-based models are in contrast with other assimilation models that have been proposed based on global misestimation of length (Brigell, Uhlarik, & Goldhorn, 1977; Jaeger & Long, 2009) and spatial averaging within attentional fields (e.g., Pressey, 1988; Pressey & Murray, 1976). We favor the edge-based approach because the CSI depends on the number of inducers (one-sided versus two-sided) and as well as the extent of the global configuration (see Fig. II.29-5; Schloss et al., 2014, Experiment 5), whereas global misestimation of length predicts that only the overall extent of the configuration should matter.

The suite of three relatively simple geometrical illusions of extent, shape, and size described here—the occlusion illusion, the CSI, and the shrinkage illusion—demonstrate rather dramatically how complex the neural processes must be that cause people to produce different judgments of corresponding properties of the same elements in the context of different spatial configurations. As Figure II.29-1 demonstrates, even two rectangles in particular configurations can give rise to measurable illusions that alter their perceived length, width, and size. At least two different mechanisms appear to be at work: partial modal completion in the occlusion illusion and edge-based assimilation in the CSI and shrinkage illusions. Moreover, these illusory influences are often mutual: what we call the inducer influences the perception of the target in the CSI, as discussed, but in most cases the target must also influence the perception of the inducer. Given the utter simplicity of the geometrical configurations involved, such illusions must be omnipresent in our everyday perceptions, even though their differences are quantitative, rather than qualitative, and relatively small. Clearly, there is much more to the perception of seemingly basic spatial properties than meets the eye.

REFERENCES

Brigell, M., Uhlarik, J., & Goldhorn, P. (1977). Contextual influences on judgments of linear extent. *Journal of Experimental Psychology: Human Perception and Performance*, *3*(1), 105–118.

Bulatov, A., Bertulis, A., Gutauskas, A., Mickiene, L., & Kadziene, G. (2010). Center-of-mass alterations and visual illusion of extent. *Biological Cybernetics*, *102*, 475–487.

Jaeger, T., & Long, S. (2009). Classical illusions from parallel line figures: Evidence for interactions among length-coding neurons. *Perceptual and Motor Skills*, *109*(2), 452–458.

Kanizsa, G. (1979). *Organization in vision: Essays on Gestalt perception*. New York: Praeger.

Kanizsa, G., & Luccio, R. (1978). *Espansione di superficie da completamento amodale*. Reports from the Institute of Psychology, University of Trieste, Italy.

Luccio, R. (1981). Effetti dimensionali del completamento amodale: un'analisi in termini di assimilazione. *Giornale Italiano di Psicologia, 8*, 375–392.

Micali, G., Giurissevich, O., & Serani, M. (1978). *Il completamento amodale come fattore dell'espansione fenomenica di superficie*. Reports of the Institute of Psychology, University of Trieste, Italy.

Morgan, M. J., Hole, G. J., & Glennerster, A. (1990). Biases and sensitivities in geometrical illusions. *Vision Research*, *30*, 11, 1793–1810.

Palmer, S. E., Brooks, J. L., & Lai, K. S. (2007). The occlusion illusion: Partial modal completion or apparent distance? *Perception*, *36*, 650–669.

Palmer, S. E., & Schloss, K. B. (2009). Stereoscopic depth and the occlusion illusion. *Attention*, *Perception*, *& Psychophysics*, *71*, 1083–1094.

Pressey, A. W. (1988). In defense of the assimilation theory of visual illusions. *Perceptual and Motor Skills*, *66*, 760–762.

Pressey, A. W., & Murray, R. (1976). Further developments in the assimilation theory of geometric illusions: The adjacency principle. *Perception & Psychophysics*, *19*(6), 536–544.

Pressey, A. W. & Wilson, A. E. (1980). Assimilation theory and the Baldwin illusion. *Italian Journal of Psychology*, *8*, 65–74.

Prinzmetal, W. (2005). Location perception: The X-Files parable. *Perception & Psychophysics*, *67*(1), 48–71.

Scherzer, T. R., & Ekroll, V. (2015). Partial modal completion under occlusion: What do modal and amodal percepts represent?. *Journal of vision*, *15*(1), 22.

Schloss, K. B., Fortenbaugh, F. C., & Palmer, S. E. (2014). The configural shape illusion. *Journal of Vision*, *14*(8), 1–18.

Vezzani, S. (1999). Shrinkage and expansion by amodal completion: A critical review. *Perception*, *28*, 935–947.

Zanforlin, M. (1981). *Perception of apparent width using adjacent surfaces: II. Apparent shrinkage and expansion of adjacent (or striped) surfaces: An explanatory hypothesis*. Report No. 37. Padova, Italy: Institute of Psychology, University of Padova.

Chapter 30
Reverse-Perspective Art and Objects—Illusions in Depth and Motion

Thomas V. Papathomas

INTRODUCTION: HISTORY

There is an interesting class of objects that exhibit a *depth inversion illusion*: namely, a point A on the object that is further from the viewer than another point B on the same object is in fact perceived to be closer than B. A corollary of the depth inversion is that these objects appear to move for a viewer who moves in front of them. In addition, when we rotate such an object about an axis, it is perceived to rotate in the opposite direction to that of the physical rotation. The hollow-mask illusion (see chapter 87 in this volume) is one of the best-known members of the family of objects that exhibit depth inversion. One of the earliest references to such phenomena is found in Ptolemy (Lejeune, 1956, p. 76; Wade, 1998), who points out that the sails of a boat may appear either as convex or concave. Sperling, Gobell, and Tseng (2009) observed a similar illusion with windsurfer sails and studied the strength of the illusion as a function of context and sail orientation. A more recent—compared to Ptolemy—member of the depth-inversion illusion family is a three-dimensional (3D) skeleton version of the Necker cube (Wheatstone, 1838).[1] In this case, the bistable cube elicits two competing percepts, alternating in time between its veridical and its reverse-depth interpretations; the violation of the occlusion rule makes it more difficult to obtain the illusory percept, making the veridical percept—which contains no such violations—dominant. In this chapter, I present two classes of physical 3D objects that elicit depth inversion based on the manipulation of perspective cues, with special emphasis on reverspectives, the class that elicits the strongest and most robust depth inversion. Note that the Ames window illusion also belongs to this perspective-based depth-inversion illusion category, but a separate chapter in this volume (chapter 14) is dedicated to it because of its rich literature and the special instructional role it has played in textbooks and video demonstrations.

PATRICK HUGHES'S REVERSPECTIVES

Patrick Hughes, the ingenious British artist, is credited as the creator of a new class of depth-inverting objects that are becoming more popular among art collectors (Griggs, 2010; Hughes, 2011; Poulton, 2010; Slyce, 2011) and vision researchers alike (Cook, Hayashi, Amemiya, Suzuki, & Leumann, 2002; Hayashi, Umeda, & Cook, 2007; Papathomas & Bono, 2004; Papathomas, 2007; Rogers & Gyani, 2010; Wade & Hughes, 1999; Wagner, Ehrenstein, & Papathomas, 2008). His first reverspective was "Sticking Out Room" (1964), as reported by Slyce (2011) and shown in Figure II.30-1a. It is composed of a single truncated pyramid—more complex reverspectives contain several such pyramids, truncated or not—that juts out toward the viewer, as shown in the three orthographic projections of Figure II.30-1b.

This is one of the fundamental building blocks for reverspectives. Notice that the near, small, truncated base ABCD of the piece is closer to the viewer than the far, large base EFGH, eliciting an inverse-depth illusion. Nevertheless, the piece is collaged in such a way as to suggest that the small base ABCD is in fact further than the large base EFGH. If this illusory percept prevails, trapezoids EADH, FGCB, EFBA, and GHDC are perceived as the ceiling, floor, and left and right walls, respectively, of a room that recedes into the wall on which it is placed. The room appears to move to a viewer who moves in front of it, as long as the inverse-depth percept is obtained.

Figure II.30-2a shows a more complex reverspective comprising two such truncated pyramids. Notice that there are two dominant competing 3D interpretations of Figure II.30-2a: First, one may perceive the true physical structure, namely, two truncated pyramids—each similar to the one of Figure II.30-1—protruding from the wall; this geometry is shown schematically on Figure II.30-2b. Second, one may view the scenery with the central group of buildings jutting out and the canals receding into the distance; this is the illusory, depth-inverting percept. The illusory percept dominates over the veridical one for a surprising range of viewpoints, even under binocular viewing, as long as there is no self-occlusion. Occlusion is a powerful cue to depth. As soon as part of the object occludes another part of the object, the illusion breaks down. Absent self-occlusion, as with most depth-inverting illusions, the scenery appears to move robustly in the mind of a moving viewer. This illusory motion is the most striking property of reverspectives.

REVERSPECTIVES, CONVENTIONAL PERSPECTIVES AND PROPER PERSPECTIVES

Theoretically, any reverspective can become a "properspective" by simply using the same scenery that is painted on the *convex* 3D surfaces comprising the reverspective and

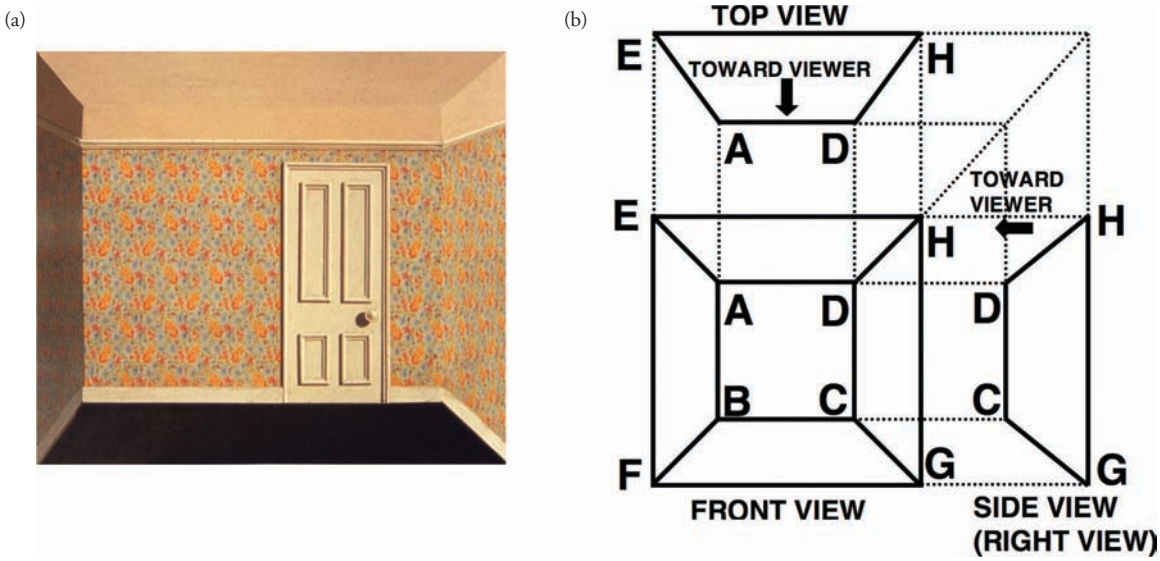


Figure II.30-1. (a) Patrick Hughes's "Sticking Out Room" (1964). (b) Orthographic projections of the piece shown in Figure II.30-1a (not drawn to scale). (Taken from Slyce [2011] with permission.)

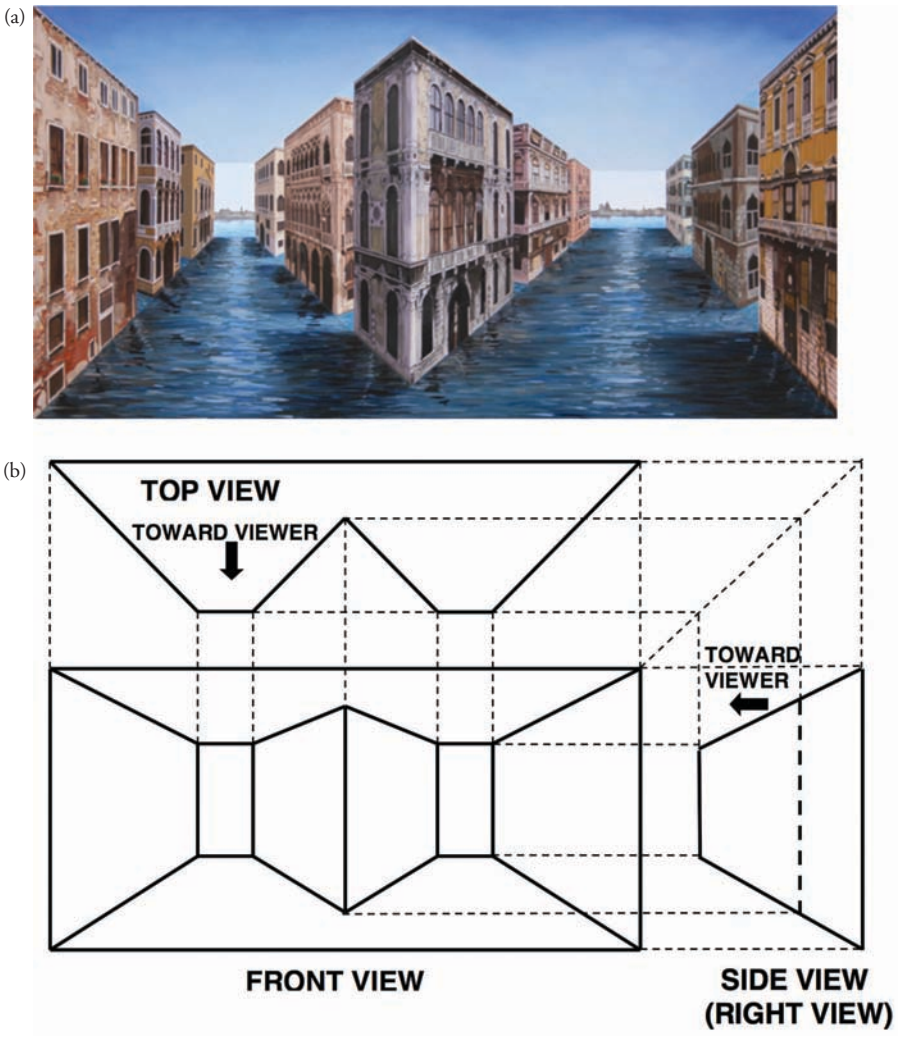


Figure II.30-2. (a) Patrick Hughes's "Vanishing Venice" (2007). (b) A schematic diagram of a reverspective with a similar 3D structure. (Taken from Hughes's website, with permission. http://www.patrickhughes.co.uk/)

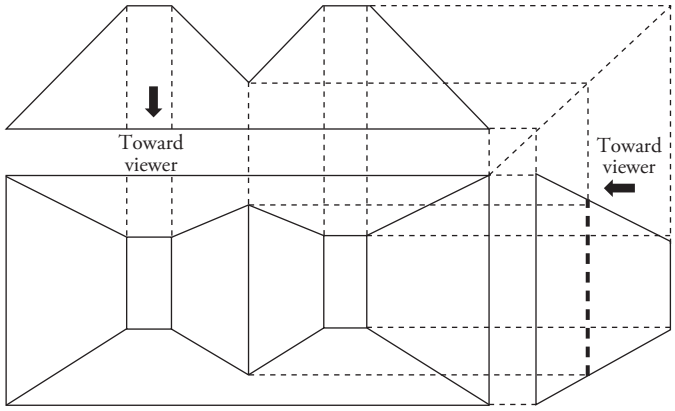


Figure II.30-3. This "properspective" (forced perspective) is the dual (hollow) structure that corresponds to the convex reverspective of Figure II.30-2b. Notice that the front views are identical.

mapping it on the reverse (*hollow)* side of the corresponding "dual" reverspective surface. For example, Figure II.30-3 shows the orthographic projections of the surface that is the dual of the surface of Figure II.30-2b. If we imagine the very same scenery of Figure II.30-2a painted on the inside hollow surface of Figure II.30-3, the result is scenery painted in *forced* perspective. Forced perspective is often used in theater stages to depict the inside of a room. The typical theater stage has a floor that is tilted down toward the audience, as well as slanted trapezoid walls with their long sides toward the audience.

When viewers see properspectives, they recover the true depth relationship of the surfaces and they have a tendency to overestimate the depth magnitude (Dobias, Moritz, Baghel, Theiler, & Papathomas, 2013); furthermore, when one moves in front of a properspective, one perceives no illusory motion. An interesting piece can be constructed using a thin sheet of material to produce the geometry of Figure II.30-2a, composed of the two convex truncated pyramids; if we view this structure from the opposite viewpoint, we now have two concave truncated pyramids that can be painted with the same scenery. An example of a two-face reverspective rotating about a vertical axis is provided on the Companion Website (see Video II.30-1).

Of course the scenery of any reverspective could also be painted on a flat canvas as a *planar perspective*. For example, the scenery of "Vanishing Venice" of Figure II.30-2a could have been rendered as a conventional painting on a flat surface, just as it appears on the flat page of this book. Planar perspectives may be perceived to move if a viewer walks past them, but the motion is weak when compared with that of reverspectives (Papathomas, Kourtzi, & Welchman, 2010). In summary, the same scenery of a reverspective can exist in at least three major forms: (a) as a reverspective that exhibits depth inversion and strong illusory motion, (b) as a forced perspective that exhibits no depth inversion and no illusory motion, and (c) as a planar perspective that elicits the percept of depth and weak illusory motion.

MORE ELABORATE REVERSPECTIVES BY HUGHES

Patrick Hughes has created a wide variety of elaborate reverspectives that go beyond the canonical structure of his earlier pieces. Of course space does not allow for a thorough treatment; interested readers can consult Slyce (2011). Here we present a few representative ingenious designs: "Reverspective Versus Perspective" (2007), "Forced into Reverse Perspective" (2008), and "Daydreaming" (2008). In the first one, Hughes combines reverspective and forced perspective on the same piece, as shown in Figure II.30-4a. The 3D piece comprises four vertical prisms that jut forward, as shown in the schematic of Figure II.30-4b.



Video II.30-1. Thomas V. Papathomas.

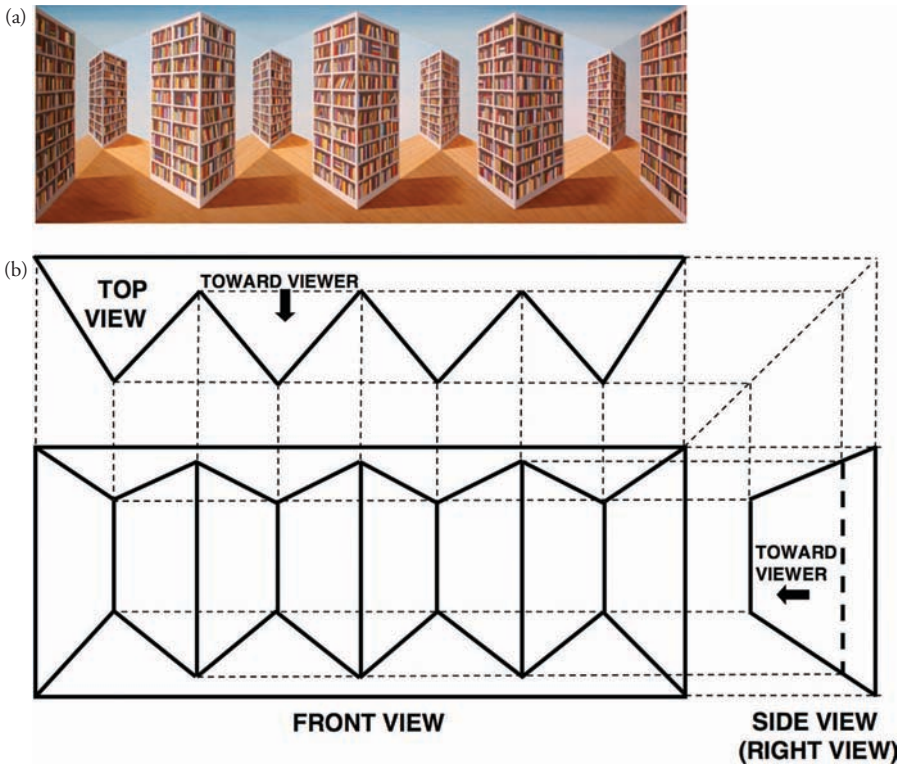


Figure II.30-4. Patrick Hughes's "Reverspective Versus Perspective" (2007). The large/small bookcases are rendered in reverse/forced perspective, respectively. (a) The painted piece. (b) Orthographic projections. (Credit: Taken from Hughes's website, with permission.)

Comparing Figures II.30-4a and II.30-4b, we notice that the large and small bookcases are rendered in reverse and forced perspective, respectively. Thus when a viewer moves in front of the piece, he or she perceives the large bookcases to turn, while the small bookcases appear stationary.

One of the most complex and ingenious pieces is "Daydreaming" (2008), depicted in Figure II.30-5a. Here, as in his earlier "Forced into Reverse Perspective" (2008), Patrick Hughes combines all three major types of perspective rendering on the same piece: reverspective, forced perspective, and planar perspective. The piece comprises five solids that jut outward: there are two truncated pyramids on either end, flanking three vertical prisms. To help understand the 3D geometry, yellow lines are used to mark the boundaries of the planar surfaces; I have also added orthographic projections in Figure II.30-5b. By inspecting Figures II.30-5a and II.30-5b, we notice that the houses in the top row/middle row are rendered in reverse/forced perspective, because they are painted on a concave/convex dihedral angle, respectively. The two middle houses in the bottom row are also in reverse perspective. Interestingly, the little houses on either end of the bottom row are painted on a planar surface. As a result, the houses that are painted in forced perspective appear not to turn for a viewer who moves in front of the art; by contrast, the "reverspective" houses appear to turn, while the planar-perspective houses appear to turn but not as strongly as the reverspective houses. Animations with "Forced into Reverse Perspective," which shares a similar 3D structure as the middle part of Figure II.30-5a containing the homes, are provided on the Companion Website (see Video II.30-2).

What about the doors (doors are one of the favorite themes in Hughes's reverspectives)? Notice that both doors are painted in reverse perspective, with their long vertical edge being further away than their short vertical edge. Thus from a straight-ahead position, the long edges—the ones with the doorknob—appear to be closer than the short (hinged) end, resulting in depth inversion. In fact, the doors form Ames windows, and they will appear to move to a viewer who moves in front of them. For example, if viewers move to the right of the straight-ahead position, they will see more of the left door. This is contrary to what they expect if they maintain the inverted-depth percept. The only logical explanation that resolves this contradiction is that the door is opening. A similar argument reveals that the right door will appear to be closing. It is confusing—and exciting—to witness this motion pandemonium!

TERMESPHERES

Artist Dick Termes invented another class of depth-inverting objects, typically in the form of six-point perspective paintings (Termes, 1998), on the surface of large spheres (usually with a radius of about 30 cm), known as "Termespheres" (Papathomas, 2007; Termes, 1971). A simple way of explaining how Termespheres are painted is to think of the artist's eye positioned at the center of a transparent infinitesimally thin spherical shell made of glass—even though this is not physically possible—and painting on this shell each point of the external world, as its "light ray" pierces the glass surface to reach the eye. Three views of two Termespheres are shown in Figure II.30-6.



Video II.30-2. Thomas V. Papathomas.

The painting is, of course, on the outside surface of the sphere. However, when observers view it for a long time, they perceive the sphere as concave, especially when the painted scene suggests a concave surface. Look, for example, at the view of the "Pantheon" in the upper left corner of Figure II.30-6. It is a bistable stimulus: it can be viewed either as a convex sphere with the apex of the dome aiming at a point above the viewer's head or as the concave part of a sphere—cut along one of its major circles—with the apex pointing up and away from the viewer. Similar depth inversions occur for the other two views of "Pantheon," as well as for all three views of "Reflecting-Through."

Most usually, Termespheres are hung from a ceiling motor and set to turn at a rate of about three to four revolutions per minute. This gives rise to a dynamic bistable stimulus. Observers perceive either a convex sphere rotating in the direction of the physical rotation or a concave sphere rotating in the opposite direction. In a similar fashion, a viewer who moves in front of a stationary Termesphere and obtains a depth-inversion illusion perceives a hollow spherical scene that appears to move.

EXPLANATIONS

The explanation that was offered for the apparent illusory motion in the hollow-mask illusion is easily extended in the case of other depth-inverting objects (Papathomas, 2007). Figures II.30-7a and II.30-7b illustrate how this explanation can be applied to reverspectives and Termespheres, respectively.

Figure II.30-7a is the top view of a reverspective (the grey region shows the physical volume covered by the art piece), seen by an observer who starts from eye position 1 (E_1), indicated by dashed lines of sight. Each point on the physical surface, such as F, can be perceived anywhere along its line of sight E_1F. Given the perceived illusory (inverse-depth) 3D scene (shown by the "peppered texture"), point F will be misassigned to position F_1, beyond its physical location. As the viewer moves to position 2 (E_2), point F can now be perceived anywhere along its *new* line of sight E_2F, indicated by a dashed line. If the viewer maintains the illusory 3D scene, point F will now be misassigned to position F_2. As shown in Figure II.30-7a, this depth misassignment

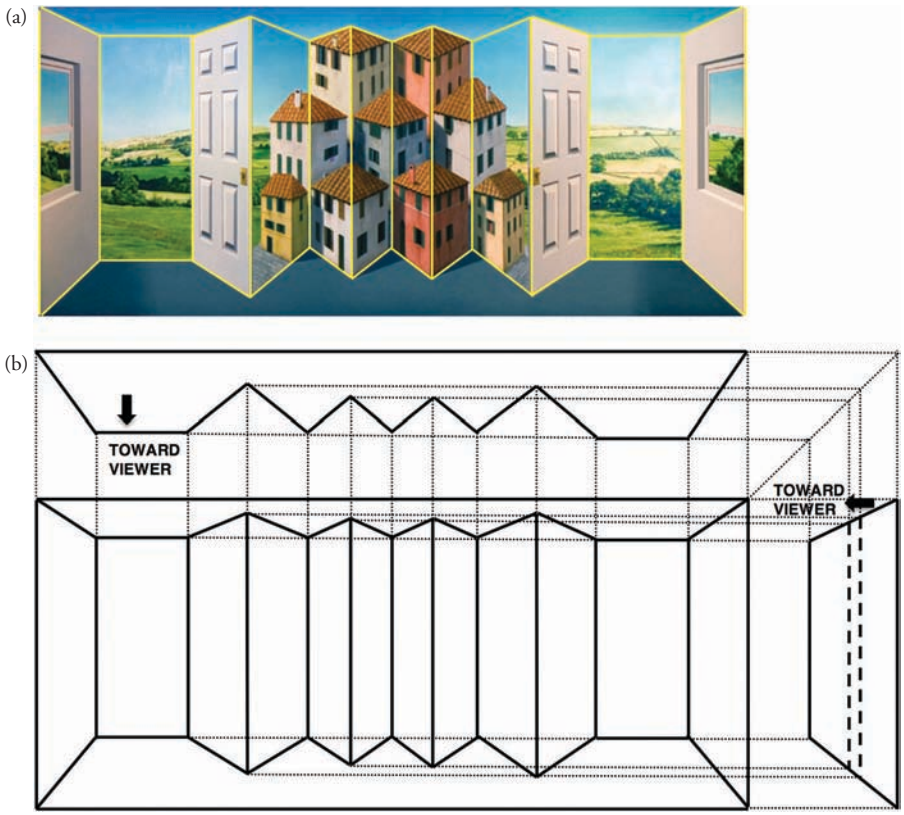


Figure II.30-5. Patrick Hughes's "Daydreaming" (2008). (a) The painted piece, with yellow lines denoting the surface boundaries. (b) Orthographic projections. (Credit: Taken from Hughes's website, with permission.)



Figure II.30-6. Three views of two Termespheres. Upper panel: "Pantheon" (1998); lower panel: "Reflecting-Through" (2006). (Credit: Taken from Termes's website, with permission. http://termespheres.com/)

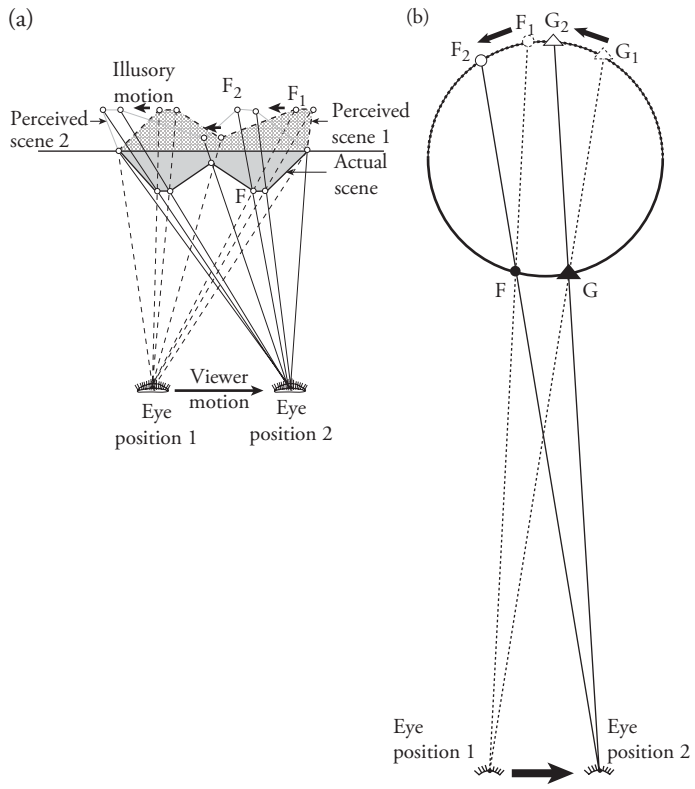


Figure II.30-7. Top views serving for a possible explanation of the illusory motion. (a) Reverse perspective. (b) Termesphere. (Modified from Figure 3 in Papathomas, T. V. (2007). Art pieces that "move" in our minds—An explanation of illusory motion based on depth reversal. *Spatial Vision, 21*, 79–95. © Brill Publishers.)

causes the illusory percept of motion: point F appears to move from position F_1 to position F_2. This occurs with every point of the reverspective, causing a compelling percept of (illusory) motion.

In a similar fashion, Figure II.30-7b presents a possible explanation for the illusory motion experienced by a viewer who moves in front of a Termesphere. The visible (front) part of the sphere is shown by a thick solid line, with two typical points F and G shown by a solid circle and triangle, respectively. The misperceived (illusory) hollow surface is shown by a thick dashed line. Points F and G are misperceived to lie on a concave surface at positions F_1 and G_1 for the starting viewer position E_1 (see dashed lines of sight). As the viewer moves over to E_2, the perceived position of physical point F has to satisfy two constrains: (a) it must lie along the *new* line of sight E_2F, and (b) if the spatial illusion still holds, it also has to lie on the illusory hollow surface (thick dashed line). These two constraints are satisfied by position F_2 on Figure II.30-7b, which is the intersection of the two constraint lines; the same goes for physical point G and its perceived position G_2. As seen on Figure II.30-7b, it is obvious from these considerations that points on the sphere will appear to move in a direction opposite to that of the viewer's direction of motion.

Cook et al. (2002), among others, formulated another plausible explanation for the illusory motion that, at first sight, appears to be different from the one presented earlier in reference to Figure II.30-7a, but it starts from the same premise: namely, that viewers build a 3D representation of the scene and they perceive illusory motion based on expectations. To illustrate, when viewers misperceive the "Sticking Out Room" of Figure II.30-1a as a concave room, they expect to see more of the right wall and less of the left wall as they move to the left. However, they get exactly the opposite results; they see more of the left wall and less of the right wall. There are two possible explanations for this unexpected outcome, both along the lines of Rock's (1983) "logic of perception": (a) their visual system realizes that its original 3D representation of a *concave* room was false and breaks the illusion, or (b) their visual system maintains the illusion of a concave room and it "reasons" that the room must be moving so as to produce the unexpected outcome. A third potential explanation is offered by Rogers and Gyani (2010): "What we see in reverspectives is the only real-world scenario (distal stimulus) that could have created the perspective and parallax transformations (proximal stimulus) that stimulate our visual systems." Their explanation has the advantage of simplicity in that it does not invoke higher-level processes or the need to rely on expectations based on previous experience.

SIGNIFICANCE

Many other 3D illusions, such as the Ames room (Dorward & Day, 1997; Gregory, 1966; Rogers, 2010; Runeson, 1988) and the Ames chair (Behrens, 1999; Gregory, 2007; Ittelson, 1952), face a common problem. The problem is that they are designed to be viewed monocularly from a peephole, thus depriving the important data-driven signals of motion parallax and stereoscopic disparity (Behrens, 1999). Perspective illusions, particularly reverspectives, solve this problem to a large extent. It is remarkable that reverspectives elicit the illusion for a moving observer who views them with both eyes open, even at small viewing distances. In this sense, they are extremely well designed to fool the human visual system; Hughes shows once more that "when it comes to understanding the brain's visual system, artists are way ahead of neuroscientists" (Griggs, 2010).

An important advantage of perspective-related illusions is that, if the hypothesis of top-down influences is valid, they offer a valuable tool as a stimulus for experiments that study the interaction of schema-driven influences (experience, memory, exposure) and data-driven signals (motion parallax, stereoscopic disparity, vergence, accommodation) in processing the visual input to arrive at a final percept. Beyond this, these illusions can be used for studying the essential problem of how the brain perceives a stable environment as we move about. Recovering true depth relationships and surface orientations under self-motion is a prerequisite for navigating successfully in the environment. Specifically, we must be able to distinguish between stationary and moving objects under self-motion. We take this ability for granted, but this is a hard problem to solve. Reverse perspectives provide powerful tools in this research because they are among rare 3D stimuli that appear to move for a moving viewer, even though they are stationary.

ACKNOWLEDGMENTS

Thanks to Steve Silverstein, Tom Grace, Marcel de Heer, Brian Keane, Michael Cooper, Akos Fehér, Anshul Jain, Rob Kohr, Aleksandra Sherman, Alexander Papathomas, Hristiyan Kourtev, Xiaohua Zhuang, Elizabeth Torres, Jillian Nguyen, Lisa Bono, Anna Zalokostas, Vanja Vlajnic, James Hughes, Jordan Ash, Michael Marks, Yushi Wang, Deepthi Mikkilineni, Arielle Yeshua, Jay Ravaliya, Nicholas Baker, Andrew Ng, Mark Theiler, Daniel Moritz, Geetika Baghel, and Joshua Dobias for their help in building reverspectives, conducting experiments with schizophrenia patients and controls, and producing demonstrations and videos. I would also like to thank Prof. Nicholas Wade, who provided me with references on the history of perspective-related illusions. Thanks to Dick Termes for providing images of his art pieces. Last but most, I am deeply indebted to Patrick Hughes for his generosity, friendship, and advice.

NOTE

1. Interested readers can consult chapter 113 in this volume. See also Wade (1998) for a comprehensive history; Wade, Campbell, Ross, & Lingelbach (2010) present a brief summary of relevant studies.

REFERENCES

Behrens, R. R. (1999). Adelbert Ames, Fritz Heider and the chair demonstration. *Gestalt Theory*, *21*(3), 184–190.

Cook, N. D., Hayashi, T., Amemiya, T., Suzuki, K., & Leumann, L. (2002). Effects of visual-field inversions on the reverse-perspective illusion. *Perception*, *31*, 1147–1151.

Dobias, J. J., Moritz, D., Baghel, G., Theiler, M., & Papathomas, T.V. (2013). Perceiving depth magnitude in reverse, flat, and forced perspective stimuli. Paper presented at the Vision Sciences Society Conference, Naples, Florida.

Dorward, F. M., & Day R. H. (1997). Loss of 3-D shape constancy in interior spaces: The basis of the Ames-room illusion. *Perception*, *26*(6), 707–718.

Gregory R. L. (1966). *Eye and brain*. London: Weidenfeld & Nicolson.

Gregory R. L. (2007). Objects and pictures. In F. A. J. L. James (Ed.), *Christmas at the Royal Institution: An anthology of lectures by M. Faraday, J. Tyndall, R. S. Ball, S. P. Thompson, E. R. Lankester, W. H. Bragg, W. L. Bragg, R. L. Gregory, and I. Stewart* (p. 315). London: World Scientific.

Griggs, J. (2010). Windows to the mind. *New Scientist*, *2778*, 34–39.

Hayashi, T., Umeda, C., & Cook, N. D. (2007). An fMRI study of the reverse-perspective illusion. *Perception*, *1163*, 72–78.

Hughes, P. (2011). *Paradoxymoron*. London: Reverspective. pp. 231–232.

Ittelson, W. H. (1952). *Ames demonstrations in perception: A guide to their construction and use*. Princeton, NJ: Princeton University Press.

Lejeune, A. (1956). *L'Optique de Claude Ptolémeé* (2d ed.). Leiden, The Netherlands: Brill.

Papathomas, T. V. (2007). Art pieces that "move" in our minds—An explanation of illusory motion based on depth reversal. *Spatial Vision*, *21*, 79–95.

Papathomas, T. V., & Bono, L. (2004). Experiments with a hollow mask and a reverspective: Top-down influences in the inversion effect for 3-D stimuli. *Perception*, *33*, 1129–1138.

Papathomas, T. V., Kourtzi, Z., & Welchman, A. E. (2010). Perspective-based illusory movement in a flat billboard—An explanation. *Perception*, *39*(8), 1086–1093.

Poulton, G. (2010, January). The art of illusion. *Centurion Magazine*, pp. 37–40.

Rogers, B. (2010). Stimuli, information, and the concept of illusion. *Perception*, *39*, 285–288.

Rock, I. (1983). *The logic of perception*. Cambridge MA: MIT Press.

Rogers B. J., & Gyani, A. (2010). Binocular disparities, motion parallax, and geometric perspective in Patrick Hughes's "Reverspectives": Theoretical analysis and empirical findings. *Perception*, *39*, 330–348.

Runeson, S. (1988). The distorted room illusion, equivalent configurations, and the specificity of static optic arrays. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 295–304.

Slyce, J. (2011). *Patrick Hughes: Perspective* (3d ed.). London: Momentum.

Sperling, G., Gobell, J., & Tseng, C.-H. (2009). Trapezoidal illusions: Windsurfers versus runways [Abstract]. *Journal of Vision*, *9*(8), 691.

Termes, R. (1971). Termes illusion or hole sphere illusion (Thesis). Otis Art Institute, Los Angeles, CA.

Termes, R. (1998). *New perspective systems: Seeing the total picture*. Spearfish, SD: Termes.

Wade, N. J. (1998). *A natural history of vision*. Cambridge MA: MIT Press.

Wade, N. J., Campbell, R. N., Ross, H. E., & Lingelbach, B. (2010). Necker in Scotch perspective [Guest Editorial Essay]. *Perception*, *39*, 1–4.

Wade, N. J., & Hughes, P. (1999). Fooling the eyes: Trompe l'oeil and reverse perspective. *Perception*, *28*, 1115–1119.

Wagner, M., Ehrenstein, W. H., & Papathomas, T.V. (2008). Vergence in reverspective: Percept-driven versus data-driven eye movement control. *Neuroscience Letters*, *449*, 142–146.

Wheatstone, C. (1838). Contributions to the physiology of vision: Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London*, *128*, 371–394.

Chapter 31
The *New* Moon Illusion

Brian Rogers and Stuart Anstis

THE ILLUSION

Consider the situation in which the sun is low in the sky and about to set and the moon has already risen 20˚ or 30˚ above the horizon. As a consequence, the elevation of the moon is considerably greater than that of the sun. In addition, choose a time in the lunar month—the first or the last quarter—when the moon is illuminated over roughly half its surface (when viewed from the earth) so that the boundary between the lit side and the dark side of the moon (known as the terminator) is almost straight. Simple physics tells us that the sun must be in a direction perpendicular to the terminator. Therefore, with the moon higher in the sky than the sun, we might expect that the terminator line be tilted *downward*[1] (forward) toward the sun. In fact, readers may wish to confirm our own observations (first made in 1980 while walking along Moon Drive in Toronto) that the terminator will invariably be tilted *backward* toward a point higher in the sky (Fig. II.31-1).

This is the *new* moon illusion, in which the sun does not appear to be in a direction perpendicular to the backward tilt of the terminator. The amount of backward tilt of the terminator increases with the subsequent phases of the moon from the first quarter until it approaches a full moon. We have called this effect, which has been noted previously but without a complete explanation (Anstis, 1991; Minnaert, 1940; Perelman, 1958; Schölkopf, 1998; Walker, 1975), the *new* moon illusion,[2] in order to distinguish it from the more familiar moon illusion studied by Kaufman and Rock (1962), Ross and Plug (2002), and others. Minnaert (1940) wrote

> immediately related to this is the observation that the line connecting the horns of the moon, between its first quarter and full moon, for instance, does not appear to be at all perpendicular to the direction from sun to moon; we apparently think of this direction as being a curved line.

How can the effect be explained? We can rule out any explanation based on the refraction of the sun's rays as they pass through the atmosphere, because this causes the sun to appear higher in the sky than it actually is, which is in the *opposite* direction from the effect described here. In fact, we can rule out any explanation based on the physics of light, including phenomena such as reflection, refraction, diffraction, and dispersion, because the sun does actually lie in a direction that is perpendicular to the terminator shadow on the moon.[3]

A little thought shows why this must be the case. As three points in space, the sun, the moon, and the observer must all lie in a single plane (Fig. II.31-2). Light from the sun to the moon must travel within this plane. Consequently, the angle of the terminator will always be at 90˚ to the plane joining the sun, moon, and observer. Moreover, it makes no difference to the angle of the terminator if the observer misjudges the distance to the sun (see later discussion) because that cannot alter the plane defined by the sun, the moon, and observer. The relative position of the observer will affect *only* the amount of the lit side of the moon that can be seen, which will vary from 100% of the surface illuminated (i.e., a full moon), when the observer is in the same direction as the sun, to 0% of the surface illuminated (i.e., a new moon), when the observer is in the opposite direction to the sun.

Readers can verify that there is indeed no physical basis for the illusion by holding up a length of string that "joins" the sun and moon together. The string from the sun to the moon will meet the terminator perpendicularly. Minnaert (1940) made the same point when he suggested "Fix this direction [sun to moon] by stretching a piece of string taut in front of your eye; however unlikely it may have seemed at first, you will now perceive that the condition of perpendicularity is satisfied" (p. 152). The illusion returns, however, when the string is taken away—the sun no longer appears to be in a direction perpendicular to the backward-tilting terminator. Not only does this demonstration prove that there is no physical basis to the *new* moon illusion, it also proves that it does not matter if the observer misjudges the distance to the sun. Joining the sun and the moon with a piece of string shows that the light from the sun does indeed travel from where the sun is perceived to be—that is, on the horizon (in this particular case). The *new* moon illusion, we argue, is a consequence of both perceptual and cognitive factors.

THE PHASES OF THE MOON

Figure II.31-3 shows how the moon will appear on subsequent days of the moon's cycle when the sun is setting on

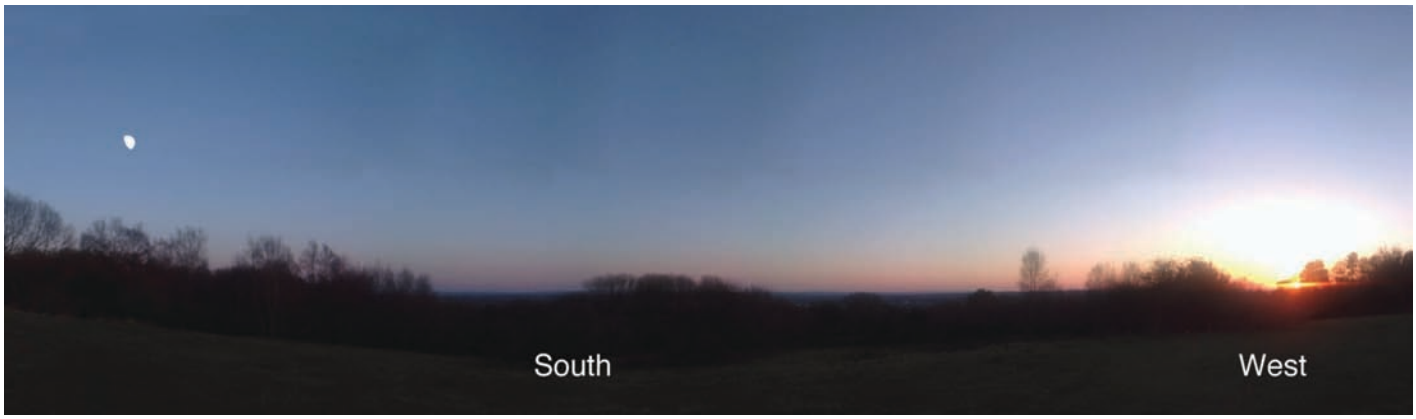


Figure II.31-1. A panoramic photograph with the sun setting in the west and the moon elevated above the horizon showing the *new* moon illusion. Note that although the moon is higher in the sky, the terminator is tilted *backward* with respect to the vertical. (Image by Brian Rogers and Stuart Anstis.)

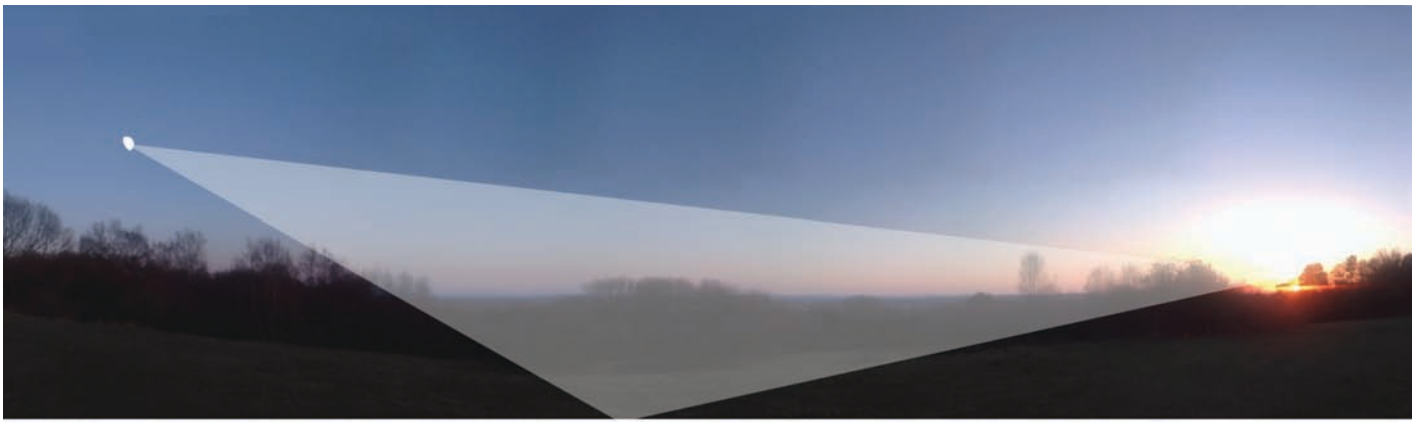


Figure II.31-2. The three points in space—the sun, the moon, and the observer—define a plane. The light that travels from the sun to the moon must travel within that plane. (Image by Brian Rogers and Stuart Anstis.)

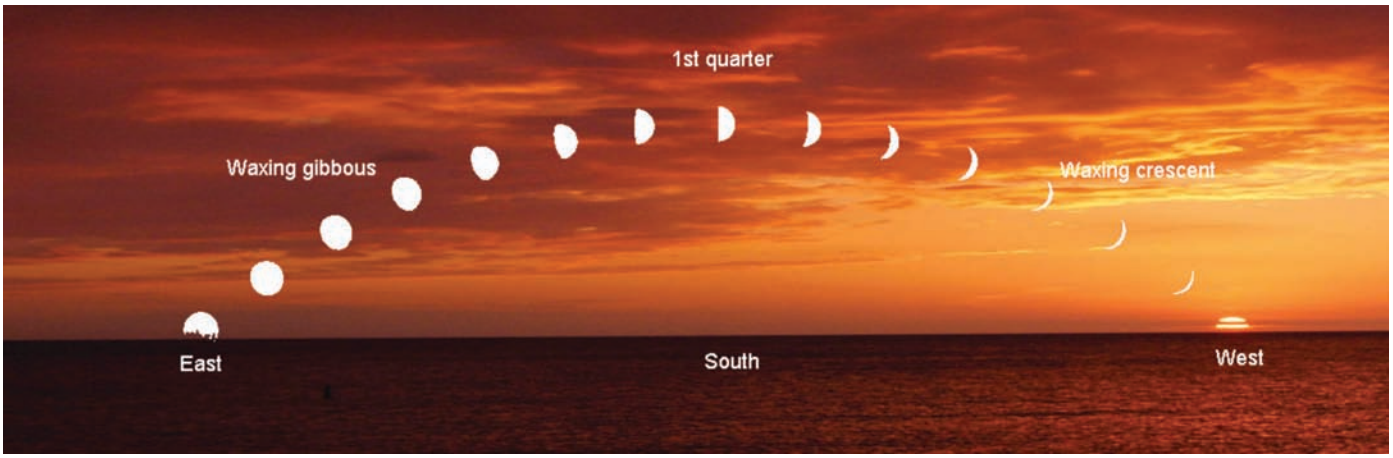


Figure II.31-3. The phases of the moon on successive days during the moon's cycle in the northern hemisphere. The right-hand half of the figure shows that, in the seven days between the new moon and the first quarter, the terminator of the moon tilts *forward*. In the seven days between the first quarter and the full moon, the terminator tilts *backward*, creating the *new* moon illusion. In all cases, the sun is just setting on the horizon in the west. The actual elevation of the moon's path depends on the season and, most crucially, on the observer's latitude. At the equator (at the time of an equinox) the forward tilt of the terminator is such that it is close to horizontal (+90˚) on the first day of the new moon and tilted backward and close to the horizontal (–90˚) on the last day before the full moon. Moving toward either the north or south poles reduces the elevation of the moon's path across the sky so that the change of tilt angle of the terminator is much smaller and stays close to the vertical throughout the moon's cycle. (Image by Brian Rogers and Stuart Anstis.)

the horizon in the west.[4] Just after a new moon, only a slim crescent is visible (the waxing crescent), and the angle of the terminator[5] is tilted clockwise[6] of vertical (i.e., tilted toward the ground) at an angle that is dependent on the elevation of the moon and the observer's latitude. As days pass and more of the moon's surface becomes illuminated, the forward tilt of the terminator is lessened and eventually becomes vertical at the first quarter, after approximately seven days. With a further passage of time, even more of the moon's surface becomes illuminated and the terminator begins to tilt *backward*, producing the ideal conditions for observing the *new* moon illusion. The terminator continues to be tilted in a more backward direction with the approach of the full moon, in the phase referred to as the waxing gibbous. Near to the equator, the tilt of the terminator can change from close to +90˚ (forward) shortly after the new moon to close to –90˚ (backward) shortly before the full moon over a period of about 14 days.[7]

Note that while the *new* moon illusion is most striking as more of the moon's surface is illuminated, it is still very apparent at the first and last quarters when the sun is setting and the terminator is vertical (Fig. II.31-3). Under these circumstances, it is difficult to imagine that a line orthogonal to the terminator—that is, parallel to the horizon—would ever intersect the setting sun that is located on the horizon. It should also be added that the illusion is not just apparent when the sun is setting at the horizon. Geometry shows (see next section) that when the sun and the moon are at a similar elevation, the moon's terminator is *always* tilted backward, rather than vertical as might have been expected, and, once again, most observers think that under these circumstances the sun should be higher in the sky rather than at the same elevation.[8]

THE GEOMETRY OF THE *NEW* MOON ILLUSION

The magnitude of the *new* moon illusion can be calculated using simple projective geometry. When the sun is setting on the horizon and phi (ø) represents the difference in azimuth angle (in degrees) between the sun and the moon and beta (ß) is the elevation of the moon above the horizon (in degrees), the angle delta (∂) (in degrees) by which the terminator is tilted[9] with respect to the vertical is given by the following simple expression:[10]

$$\partial = \arctan[\cot(\varnothing)\sin(\beta)] \qquad \text{Equation 1}$$

This expression shows that both ß and ø affect the tilt of the terminator with respect to the vertical. The tilt of the terminator increases with an increase in the elevation of the moon (ß) and decreases (i.e., becomes closer to vertical) as ø approaches 90˚. As the azimuth angle increases beyond 90˚ (after the first quarter), the sign of the tilt switches from being positive (i.e., tilted forward toward the horizon) to negative (i.e., tilted backward). This is when the *new* moon illusion is most obvious. The maximum backward tilt occurs when the azimuth angle approaches 180˚ (when the moon becomes a full moon). Note that there is no distance term in the equation. The size of the backward or forward tilt of the terminator is quite independent of the distances to the sun and moon and depends only on their elevations and the azimuth angle between the two.

The expression also shows that when the moon is at its first or last quarter (ø = 90 ˚ degrees), ∂ is zero (i.e., the terminator is vertical) irrespective of the elevation of the moon. When the azimuth angle between the sun and the moon is 135˚ (waxing gibbous) and the moon is elevated 20˚ above the horizon, there is a striking *new* moon illusion, and the expression predicts a backward tilt of the terminator of about 19˚ degrees.

It is also interesting to see what happens when the sun and moon are at the same angle of elevation in the sky. Under these conditions, the tilt of the terminator (∂) is given by an equally simple expression:

$$\partial = -\arctan[\tan(\varnothing/2)\sin(\beta)]^{10} \qquad \text{Equation 2}$$

The negative sign shows that ∂ is always negative, and consequently the terminator will always be tilted *backward* when the elevations of the sun and the moon are the same, whatever the azimuth angle between the sun and moon. Once again, the extent of the backward tilt increases with the common elevation (ß) of the sun and moon. For example, when the sun and moon are both elevated 20˚ and the azimuth angle between the sun and the moon is 90˚, the expression predicts a backward tilt of about 19˚. Finally, the expression shows that the backward tilt increases with an increase in the azimuth angle between the sun and the moon. For example, for the same elevation of the sun and moon of 20˚, an increase in the azimuth angle from 90˚ to 135˚ increases the backward tilt from about 19˚ to about 40˚.

CAUSES OF THE ILLUSION

A Cognitive Component

So much for the geometry. What is the cause of this very striking effect? First, there is what can be referred to as a "cognitive" component. Most people make the assumption that if the terminator is tilted backward (with respect to the vertical), the sun should be higher in the sky than the moon. This sounds reasonable because if the terminator is tilted backward, a line orthogonal to the terminator will point toward a location higher in the sky. It turns out that this assumption is unsound if the angular separation between the sun and moon is more than a few degrees (see later discussion). Note that relying on this assumption is a cognitive rather than a perceptual mistake, because we assume that the sun should be higher in the sky even if the sun is not visible. However, even if we did not make this assumption, there is still a strong illusion that an imaginary line joining the sun and the moon would not intersect the terminator orthogonally. This is a purely perceptual effect that also needs to be explained.

A Perceptual Component

We believe that the principal cause of the perceptual component of the illusion is perspective—that lines that are physically straight and parallel in the world are not parallel in projection. The angular separation between parallel lines (the top and bottom edges of a long wall, for example) is

greatest at the point where the wall is closest to the observer and decreases on either side. This can be seen when parallel lines in the world are projected onto Helmholtz's "Blickfeld" or *celestial sphere*—the 19th-century equivalent of Gibson's optic array (Helmholtz, 1910; Gibson, 1979; Rogers & Brecher, 2007, Rogers & Rogers, 2009; Fig. II.31-4a).

The celestial sphere is a sphere with the eye at its center; the light from the surrounding world intercepts and creates an image on the surface of the sphere. The spherical shape ensures that equal distances on the celestial sphere subtend equal angles at the eye. (To a first approximation, equal angles subtended at the eye also project to equally spaced locations on the roughly spherical retinal surface.) There is, however, a difference between the properties of the celestial sphere and those of the retinal image. The retinal image, or Helmholtz's "Sehfeld," changes with movements of the eye whereas the projection of the surrounding world onto the fixed celestial sphere ("Blickfeld") remains the same.

Figure II.31-4a shows that straight lines in the world project to the *great circles* that join the "poles" on either side of Helmholtz's celestial sphere. This figure also makes explicit that the *angular separation* between parallel lines in the world is not constant but increases from one pole to a maximum before decreasing toward the other pole. In contrast, the lines in the world on a frontal surface that project onto *lines of latitude* on Helmholtz's celestial sphere (lines that have a *constant* angular separation) are diverging on a frontal surface—with a pin-cushion distortion—rather than being parallel (Fig. II.31-4b; Rogers & Brecher, 2007). Figure II.31-4 shows the geometry of the situation, but it is an empirical question, first addressed by Helmholtz, as to which lines in the world are *judged* to be straight and parallel.[11] In many cases, we are unaware of the consequences of perspective and straight lines that are physically parallel are often seen as being parallel even though they project to great circles whose angular separation is not constant (see later discussion). This may be due to size constancy processes that take account of distance. For example, we may judge that the wall in front of us to be of constant height even though the angular separation between the floor and ceiling contours is not constant. In other situations, however, such as Helmholtz's (1910) famous chessboard pattern, the contours that are seen as straight and parallel are, in reality, curved (Fig. II.31-5).

In an experiment that used a modified version of Helmholtz's chessboard pattern, Rogers and Rogers (2009) reported that an observer's choice of the lines that appear to be straight and parallel is dependent on the disparity-specified shape of the surface against which the lines are seen. The pair of lines that appear to be parallel when they are seen against a flat, frontal surface are those that are approximately parallel on that flat surface, despite their

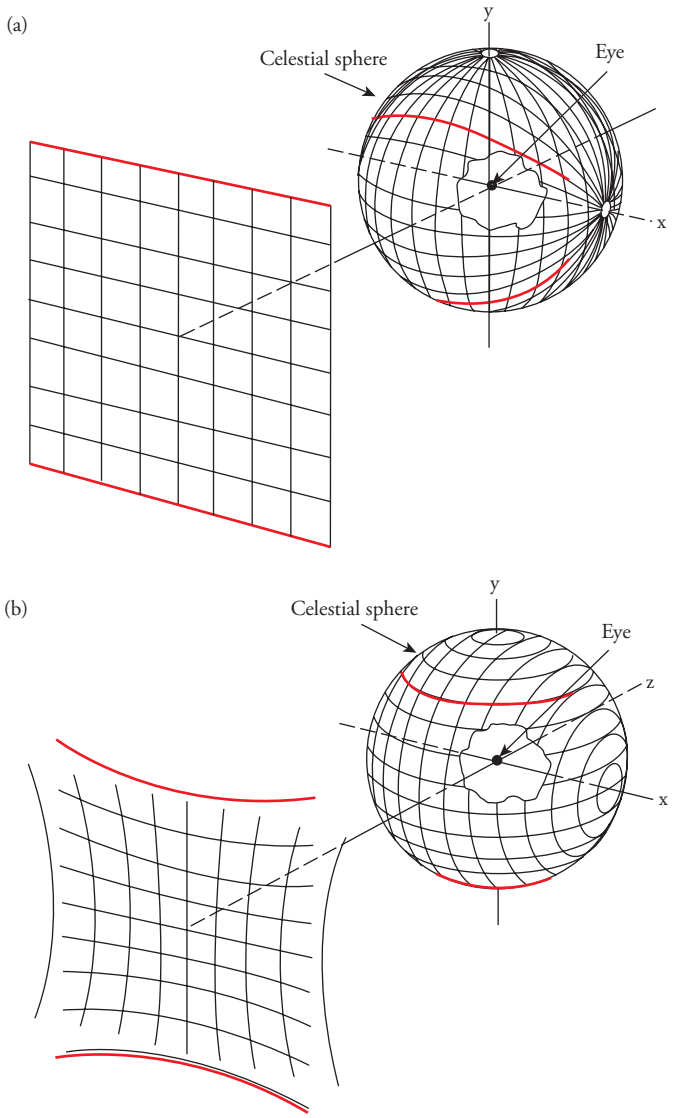


Figure II.31-4. (a) Straight and parallel lines in the world project to great circles on Helmholtz's celestial sphere. In contrast, the lines in the world that project to horizontal lines of latitude (with a constant angular separation) are diverging on a frontal surface with a pin-cushion distortion (b). (Image by Brian Rogers and Stuart Anstis.)

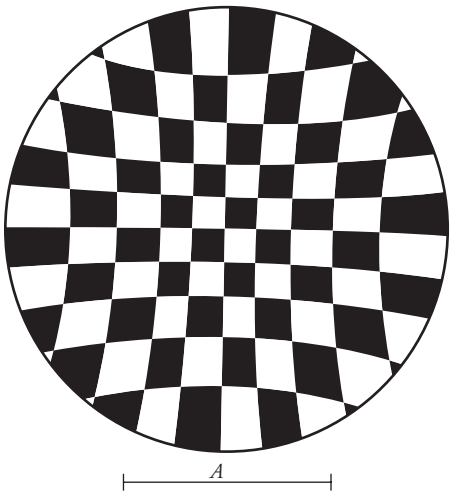


Figure II.31-5. When viewed from a very close distance (the length of the short horizontal line A), the curved contours of Helmholtz's distorted chessboard pattern appear to be straight. (Reproduced from Figure 22 of Helmholtz's "Treatise on Physiological Optics." The original book was published by Voss in 1910, as indicated in the References.)

continually changing angular separation. On the other hand, the pair of horizontal lines that appear to be parallel when they are seen against a large vertical cylindrical surface (with the eye at the center of the cylinder) are also those that are approximately parallel across the cylindrical surface (and thus of constant angular separation). Rogers and Rogers concluded that judgments about which lines are straight and parallel are crucially dependent on the surface against which the lines are seen.

In other words (and in answer to Helmholtz's question [see note 11]), there are no particular retinal arcs that represent the loci of straight and parallel lines in the world. The relevance of Rogers and Rogers' (2009) findings to our *new* moon illusion lies in the fact that both the horizon and the imaginary line joining the sun and moon are straight but the line joining the sun and moon is typically seen as curved against the dome of the sky.

In the case of the *new* moon illusion, there is obviously no disparity information about the distance of the light rays traveling from the sun to the moon. Instead, the only distance information (albeit incorrect) is provided by the flattened dome of the sky against which the lines are seen. The projections of the straight lines in space that intercept the flattened dome of the sky are similar to the great circles on a spherical projection surface (Fig. II.31-4a). Figure II.31-4a also shows that the angular separation of straight lines in the world increases from zero at one pole to a maximum before decreasing toward the other pole. The same is true of the separation between the horizon line and the straight line joining the sun to the moon. For example, when more than half of the moon's surface is illuminated (in the waxing gibbous phase) and the terminator is tilted backward, the great circle joining the setting sun to the moon (which is actually a straight line) initially increases in elevation before reaching a maximum at an azimuth 90° to the sun's direction (south at two equinoxes) and then decreases until it intersects with the moon (Fig. II.31-6).

In the most extreme case, when the moon is almost full and the observer is at the equator, the terminator will be tilted backward and close to horizontal. The great circle joining the sun and the moon will pass almost directly overhead (with an elevation close to 90°), even though both the sun and the moon are close to the horizon. In his *Flying Circus of Physics*, Walker (1975) argues that the presence of the dome of the sky is the cause of the illusion. He writes: "The illusion results from your perception of the sky as an overhead spherical dome" (p. 285). It should not surprise us that two parallel lines projected on to a spherical or near-spherical dome as great circles are perceived as diverging and then converging as they cross the dome, because if the same pair of lines were actually *drawn* on the dome's surface, rather than projected on it, they would diverge and then converge on that surface, and that would be a correct rather than an illusory percept.

In case one might think that the illusion has something to do with the fact that one cannot actually see the path of the light rays from the sun to the moon, imagine the situation where the path of those light rays is defined by the contrails of an airplane that is flying in a straight-line path across the sky in the plane defined by the observer, the sun, and the moon (Fig. II.31-2). In this case, the contrail will still appear to be curved with respect to the horizon line. Minnaert (1940) made a similar point when thinking of a searchlight beam directed across the sky:

> A searchlight casts a slender beam of light horizontally over a wide open space. Although I know that the beam runs in a perfectly straight line, I cannot get rid of the illusion that it is curved, highest of all in the middle and sloping down to the ground on both sides. (p. 151)

However, the fact that we see the sky as a flattened dome does not explain why the line joining the sun and the moon appears to be curved while the horizon line (which is just another great circle) appears to be straight. If the effect were solely due to the flattened dome of the sky, it should be possible to see the line joining the sun and moon as straight and the horizon line as concave upward (U-shaped), but this is never the case. The role of the flattened dome in our *new* moon illusion could be tested, in principle, by taking away the dome and imagining the situation of an astronaut looking out of the spaceship and being able to see the parallel rays of light emanating from the sun or a distant star. Those crepuscular rays, although physically parallel, diverge in projection as they leave the sun or star, then travel along the great circles and ultimately converge at a point in the opposite direction to the sun or star. While any individual ray in isolation might be perceived as straight, visible rays of either side are likely to appear as curved

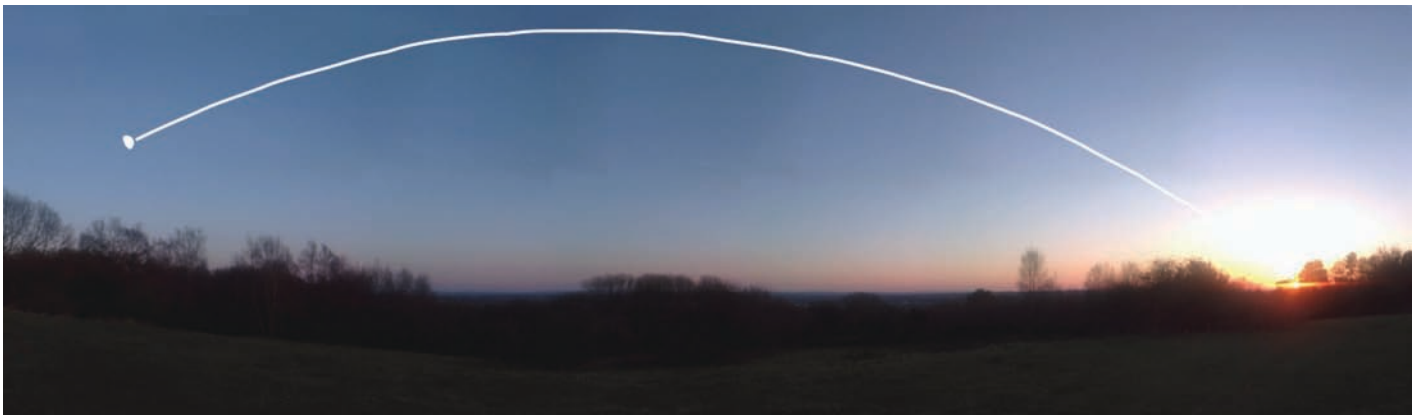


Figure II.31-6. A depiction of the great circle (which is actually a straight line) that joins the sun to the moon. (Image by Brian Rogers and Stuart Anstis.)

with respect to the first ray because the angular separation between the rays is smaller in the direction of the star (and in the opposite direction) and larger in between. Thus it would appear that, in the *absence* of any distance information including the flattened dome of the sky, we see parallel lines as they are imaged on our retinas—lines that diverge and then converge.

Our explanation of the *new* moon illusion—that parallel lines are seen as diverging and then converging when there is no information about the shape of the surface against which the lines are seen—leads to the prediction that there should be no comparable illusion in a laboratory situation where there is plenty of information about distance and surface shape. Consider lying close to floor level and looking upward toward the horizontal boundary between the opposite, frontal wall and the ceiling. Position a flashlight in (say) the upper right-hand corner illuminating a circular ball positioned in the upper left-hand corner, both elevated by the same amount with respect to the floor. As a consequence, the light illuminating the ball must travel along the wall–ceiling boundary. It must also be true that the terminator between the lit and dark sides of the ball will be orthogonal to the line of the wall–ceiling boundary. However, because of perspective, the terminator will not be vertical (even though it is orthogonal to the line of the wall–ceiling boundary), in the same way that the terminator on the moon is not vertical when the sun and moon are at the same elevation (Equation 2). However, and unlike the *new* moon illusion, the line of the wall–ceiling boundary will not typically be seen as curved. This is because in the laboratory situation we have ample information that the wall lies in a frontal plane and thus lines that are truly straight and parallel will be seen as straight and parallel (Rogers & Rogers, 2009).

CONCLUSIONS

We conclude that the explanation of the *new* moon illusion involves two factors—one cognitive and one perceptual. The *cognitive* factor is that we wrongly assume that a backward-tilted terminator of the moon is a reliable and valid indicator that the light source creating the terminator is at a *greater* height above the ground plane. In fact, the light source could be higher, lower, or at the same height, because of the effects of perspective. This cognitive factor provides a satisfactory explanation of why we make a mistake in *predicting* the sun's location with respect to the moon, even when the sun is not visible to the observer.

There is, however, also a purely *perceptual* aspect to the illusion. The sun does not *appear* to be in a direction perpendicular to terminator, even though it actually is, and, in addition, a line joining the sun and the moon appears to be curved. As we have argued earlier, if nothing else were visible apart from the sun and the moon and a hypothetical jet contrail joining the two, it seems certain that (a) the sun would appear to lie in a direction perpendicular to the terminator and (b) the contrail would appear to be straight: in other words, there would be no illusion. Moreover, the contrail would still appear to be straight even if it were seen to lie on a vast sphere surrounding the observer, as long as there was no visible horizon, ground plane, or other reference line. Thus the presence of the flattened dome of the sky does not appear to be *necessary* as an explanation of the *new* moon illusion, although it may play some role. On the other hand, it would seem that the presence of the ground surface and visible horizon do play a role both as a comparison great circle and as a horizontal reference surface. To test this, it would be interesting to see what would happen to the illusion if the typical *new* moon illusion scenario were inverted optically so that the sun, moon, and observer lay in a plane close to the (invisible) physical horizon and the horizon and ground surface of the scene were upside down and located above. Would the contrail joining the sun and moon appear to be straight and the horizon line of the scene higher up in the field of view appear to be bowed in a downward concave curve? The experiment remains to be done.

ACKNOWLEDGMENTS

This work was originally supported by Science and Engineering Research Center Grant GR/E 9556.9 and EC Grant ESPRIT 3001 to BJR and by Grant A0260 from the Natural Science and Engineering Council of Canada to SMA.

NOTES

1. With respect to the vertical.
2. Note that it has nothing to do with the new moon, or it would be called the *new moon* illusion.
3. Reflection, refraction, diffraction, and dispersion have also been ruled out as explanations of the traditional moon illusion since the angular sizes of the horizon and zenith moons are the same.
4. The actual elevation of the moon varies with the seasons as well as the phase of the moon and is a function of the earth's tilt and the inclination of the moon's orbit around the earth with respect to the ecliptic.
5. Even though the terminator is not a straight line, apart from at the first and last quarters, we can define the terminator angle as the angle between the vertical and the chord that joins the horns of the waxing or waning crescent or the chord that joins the two points where the totally illuminated side meets the shallower curvature of the waxing or waning gibbous.
6. Counterclockwise in the southern hemisphere.
7. The cycle then reverses after the full moon—through the waning gibbous, last quarter, and waning crescent phases—before the coming of the next new moon.
8. Schölkopf (1998) measured the size of the illusion during the waxing crescent phase (before the first quarter; see Fig. II.31-3) and reported that observers' settings of the terminator of an *artificial* moon showed a constant error that averaged 12° with the terminator tilted more clockwise (i.e. more downwards) than the real moon's terminator: These results provide evidence of a *new* moon illusion in the days *before* the first quarter of the moon.
9. A +ve sign indicates a forward tilt and a –ve sign a backward tilt.
10. We are extremely grateful to Michael Berry for spotting the errors in our original formulation of Equations 1 and 2 and for providing us with the correct versions.
11. Helmholtz (1910) asked the question: "Welches sind die ungekrümmten linien im sehfelde?" [What are the uncurved lines in the field of vision?]

REFERENCES

Anstis, S. M. (1991). Hidden assumptions in seeing shape from shading and apparent motion. In A. Gorea, Y. Fregnac, Z. Kapoula, & J. Findlay (Eds.), *Representations in vision* (p. 281). Cambridge, UK: Cambridge University Press.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Helmholtz, H. von. (1910). *Treatise on physiological optics* (J. P. C. Southall, Trans.). New York: Dover, 1962.

Kaufman, L., & Rock, I. (1962). The moon illusion. *Scientific American, 136*, 953–961.

Minnaert, M. (1940). *Light and colour in the open air*. London: G. Bell and Sons.

Perelman, I. A. (1958). *Astronomy for entertainment* (A. Shkarovsky, Trans.). Moscow: Foreign Languages Publishing House.

Rogers, B. J., & Brecher, K. (2007). Straight lines, "uncurved lines," and Helmholtz's "great circles on the celestial sphere." *Perception, 36*, 1275–1289.

Rogers, B. J., & Rogers, C. (2009) Visual globes, celestial spheres, and the perception of straight and parallel lines. *Perception*, *38*, 1295–1312.

Ross, H., & Plug, C. (2002). *The mystery of the moon illusion: Exploring size perception*. Oxford: Oxford University Press.

Schölkopf, B. (1998). The moon tilt illusion. *Perception*, *27*, 1229–1232.

Walker, J. (1975). *The flying circus of physics: With answers*. New York: Wiley.

Chapter 32
Geometrical Errors Are the Cost of Maintaining the Luminance Contrast Polarity

Sergio Roncato

INTRODUCTION

Individuals who are asked to compare Figures II.32-1a and II.32-1b are likely to take note of only one common feature: a series of rectangles in different shades of grey. It is unlikely that anyone will contend that the pattern of rectangles in the first figure is the same as that in the latter, since the impression of horizontal waves elicited in the former is as vivid as the impression that the rows of rectangles are straight in the latter. To convince observers that the difference is the result of an illusory effect, we must ask them to use a ruler as a guide to align the extremities of the rectangles. Only then does it become apparent that they both lie on a straight line.

Why is the alignment of geometrical shapes affected by alterations in the background luminance? Misperceptions such as this one induced by surrounding surfaces or configurations are not a rarity and belong to the category of so-called "contextual effects," which refer to the influence a background context can have on object perception. Some well-known examples of this effect are illustrated by the Ebbinghaus illusion and the rod-and-frame illusion. In the first, the same object appears larger when surrounded by small objects than when it is surrounded by other large objects; in the second, the rotation of a frame induces an illusory slant of an inner line. In these cases a variation in a parameter in an object's surroundings affects that object's perception. There are also interesting cases in which an image appears to change because an event of a different nature. According to the Benary cross effect (Benary, 1924), the brightness of a small triangle varies in relation to its "figural belongingness": it is brighter if it is perceived as a part of a black cross and darker if it is external to the cross perimeter.

The illusion we examine in this chapter belongs to this class of "hybrid effects," as object alignment appears to be altered by changes in the background luminance. But how can a change in illumination cause such a vivid geometrical deformation?

To answer this question we need to re-examine the phenomenon known as "contour interpolation," or the ability of the human visual system to integrate fragments into a whole. Evident manifestations of the phenomenon are visible when an illusory contour appears to fill a gap between separated edges or when a figural unit emerges in the form of an aligned group out of a multitude of others. These effects are observed provided that some spatial constraints (distance, orientation) are respected. Luminance parameters may play a crucial role in some figural contexts such as in those studied by Roncato and Casco (2003) and Parlangeli and Roncato (2008), who ascribed the apparent deviation from straightness to a conjunction of closely adjoining parallel edges. A phenomenological demonstration is outlined in Figure II.32-2a.

In this figure, the lighter tiles seem to bend at the corners moving closer to the darker ones. The effect is visible, although less intensely, at the corners of the darker tiles. The illusion disappears when, as in Figure II.32-2b, the background luminance becomes the lightest surface in the configuration. The same effect can be obtained by filling the background with the darkest grey shade. That variation has important consequences on the so-called contrast polarity (CP), that is, the direction of luminance variations (from light to dark or from dark to light) across edges. In Figures II.32-2c and II.32-2d the signs "+" and "–" symbolize the lighter and darker sides, respectively, of the edges and make it possible for us to check the contrast polarities and how they change when the background goes from an intermediate grey to a lighter shade of the same color. The edges with the same CP in Figure II.32-2a have the opposite one in Figure II.32-2b.

Roncato and Casco (2003) and Parlangeli and Roncato (2008) described how, in the former case, the edges with the same CP tend to bend at the extremities due to a phenomenon of boundary completion in the presence of two conditions: (a) the extremities are in close proximity and (b) they share the same CP. The phenomenon seems then to correspond to the *CP rule*: that is, in the event there are more than one directions of contour completion, the one preserving the CP will perceptually prevail.

The rule applies when the separation between edges does not exceed a certain limit. When the separation measured is in a direction perpendicular to the edges, the maximum distance between edges appearing conjoined is about 8 arcminutes of visual angle. The range widens when the direction is longitudinal: two nonoverlapping edges tend to conjoin if the gap between the two opposite extremities does not exceed 18 arcmin of visual angle. In this situation

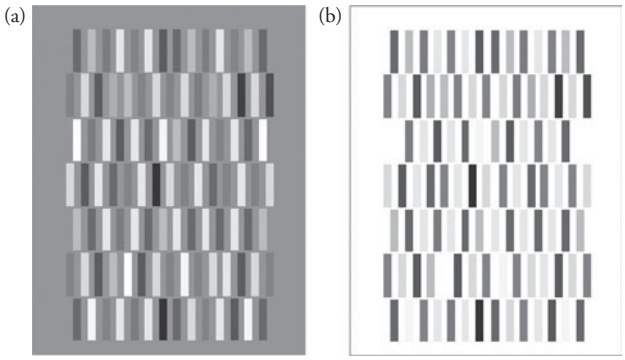


Figure II.32-1. The difference between the two figures lies in the background luminance. The pattern of rectangles in the two figures is identical as far as the color of the rectangles and the geometrical arrangement is concerned. The wavelike pattern in (a) is illusory. (Author's own unpublished work.)

the effect is short-ranged and is an extension that researchers on contour completion have agreed to assign to the CP effect (Dresp, 1999).

The conditions for the CP rule to apply are met in Figure II.32-1a. This can be verified in Figure II.32-2e where an enlarged portion of Figure II.32-1a including "+" and "–" symbols used to indicate the light and the dark sides, respectively, of the edges has been reproduced. Here

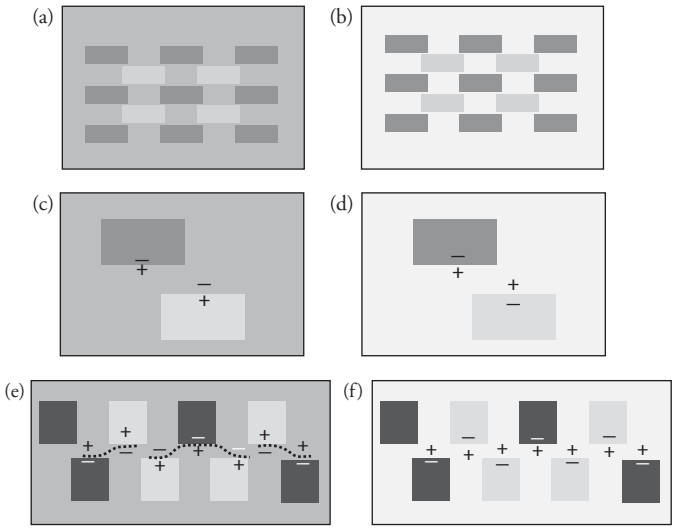


Figure II.32-2. (a) Rows of light and dark rectangles are drawn against an intermediate grey color. (b) The figure is the same as in (a); the only difference is the background, which is a lighter shade of grey. (c) An enlargement of a portion of (a). "+" stands for the lighter and "–" for the darker side of the edge. Conventionally two edges have the same CP when, following the same direction, the same signs are seen. The lower edge of the dark square on the left has the same CP ("–/+" going downward) as the upper side of the lighter square on the right. (d) The same as (c) but with a lighter background with respect to the other two. The CPs of the sides in close proximity are now opposite ("–/+" vs. "+/–"). (e) An enlargement of a portion of Figure II.32-1a with the CPs of the edges of two adjacent rows. Dotted curved lines conjoin the edges of the same CP in close proximity. This path of conjunction provides a local slant causing misperception of the square alignment. (d) Enlargement of a portion of Figure II.32-1b. Since all the edges bordering the interspaces between the two rows of squares have opposite CP, no interaction is expected. (Author's own unpublished work.)

the CP is determined following the downward direction for horizontal edges.

If we inspect the space separating the two rows of rectangles, we find that the shorter sides of the rectangles overlap slightly and are closely attached. In this situation one of two situations are possible: the edges can share the same or have opposite CPs. In the former case, the CP rule leads us to predict a contour-completion phenomenon, while in the latter no interaction is expected between the edges. The dashed curved line indicates the "path of conjunction" of edges with the same CP and highlights the illusory tilt that was predicted. In Figure II.32-2e two tilts, in the left half, are followed by two tilts in the opposite direction. Cyclically replicated, the phenomenon gives rise to a wavelike impression. The CP rule cannot be applied in Figure II.32-2f because the CPs are always opposite at the rectangle corners, and this the reason the straightness is never misperceived.

The CP rule provides an alternative explanation for the café wall illusion and for its inversion when the "mortar line" (the line interposed between two rows of tiles) is moved a short distance from the line dividing the tiles. An example of this illusion is reproduced in Figure II.32-3a: the parallel mortar lines only seemingly converge at their extremes. If the mortar line is moved, the direction of the illusion is inverted, as can be seen in Figure II.32-3b, where the horizontal lines appear to diverge at their extremities (Parlangeli & Roncato, 2008).

Enlargements of portions of these examples in Figures II.32-3c and II.32-3d illustrate the outcome of the CP rule. The interacting edges are the lower and the upper sides of the tiles of adjacent rows and the CP has been indicated for each of them. The dotted curved line indicates

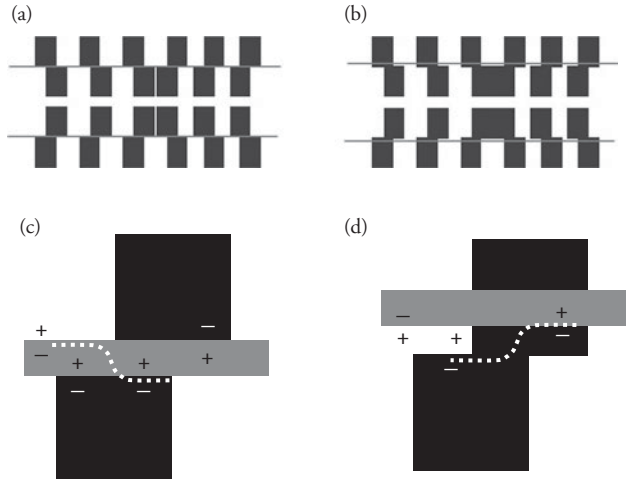


Figure II.32-3. (a) A version of the café wall illusion. While the two horizontal mortar lines are straight, they appear to converge at the extremities. (b) This is the same as (a) but with the horizontal mortal lines displaced upward or downward. Here the straight mortar lines appear to diverge at the extremities. (c) An enlargement of a portion of (a). The dotted curved line indicates the path of conjunction of edges sharing the same CP. Its inclination predicts the overall tilt of the mortar line. (d) An enlargement of a portion of (b). The dotted line indicates the path of conjunction of edges with the same CP in close proximity. In agreement with the illusory effect observed in (b), the predicted inclination is in the opposite direction with respect to (c). (Author's own unpublished work.)

the direction of the conjunction of edges with the same CP. Here the local slant predicted by the CP rule is congruent with the phenomenological tilt or illusory convergence outlined in the traditional version of the café wall illusion. Predicting what happens when the mortar line is moved is a little more complicated. The enlargement of Figure II.32-3b illustrates the pattern of boundary interaction, and the dotted lines illustrate the predicted direction of edge conjunction. Here the illusory tilt is moving in the opposite direction with respect to what is perceived in Figure II.32-3a. The CP rule is then a good predictor of the illusory tilt of the café wall illusion in its different variations.

A third example illustrates the strength of this rule and at the same time its independence from the action of other organizing factors. There are two checkerboards in Figures II.32-4a and II.32-4b. The left one seems to have been drawn in perspective, that is, with the vertical lines seeming to converge. In the right one, checkers are arranged away from the square centers, partly cover horizontal edges, and are tangential to a vertical dividing line. The checkers are isochromatic disks on the rows and are alternately dark and light disks in the columns.

This configuration was analyzed by Parlangeli and Roncato (2010) with the aim of observing boundary-completion in situations in which the CP rule competes with the Gestalt law of organization known as "good continuity." The competition arises at the convergence of four surfaces (*X*-junction) where a vertical dividing line intersects a horizontal one (Figure II.32-4c). Each rectangle side can continue straight along the collinear edge or can deviate by 90°, going around the rectangle perimeter. In any of the two cases, one rule is obeyed and the other is violated. When the rectangle side continues along the collinear side, the Gestalt law of good continuity is obeyed but at the cost of inverting the CP. When the side turns perpendicularly (following the rectangle perimeter), the CP is preserved at the cost of violating the good continuity law due to an abrupt change in direction. The regions of the checkerboard in which four surfaces converge present a condition in which the law of good continuity is challenged by another binding rule. One rule can prevail, leading to the perception of linearity, but the prevalence can be momentary and can be quickly replaced by the perception of squares or stairs.

In a condition of instability the appearance of a new path of boundary conjunction can prevail. Figure II.32-4d illustrates how this can take place. A portion of the *X*-junction is hidden by dark circles that are tangential to two rectangle sides. As a result, a new direction of boundary interaction appears because the CP is preserved at the cost of a slight deviation from straightness.

This path of conjunction is illustrated in the enlargements of Figure II.32-4d with grey S-lines. These lines bend at the two extremities following the same clockwise direction so that the whole arc–vertical edge arc pattern becomes a local slant cue. It is assumed that the effects of a series of these paths of conjunction along the same vertical axis are summed, giving rise to an overall tilting effect. In the flanking axes a symmetric pattern of luminance generates an illusory tilt in the opposite direction so that the two axes appear to converge/diverge.

The phenomenon appears to be particularly relevant, because in Figure II.32-4b it occurs in violation of the perspective checkerboard effect. In fact, instead of perceiving two converging columns of disks in the central vertical region, these seem to diverge. The illusion seems to have neutralized or hidden two important sources of information: the geometrical lengths (spatial extent) and the apparent depth cues. But how can converging and diverging directions coexist in the same image? Figure II.32-4b demonstrates that our visual system searches for a "solomonic solution" by splitting the image into two layers: one for the checkerboard surface and the other for the checkers. In this particular representation in which depth is stratified, two images coexist, but there is a curious interference. Even if the disks are all the same size, the ones closest to the top—or the region receding in depth—seem to be larger. This is the consequence of an illusory effect that takes place when

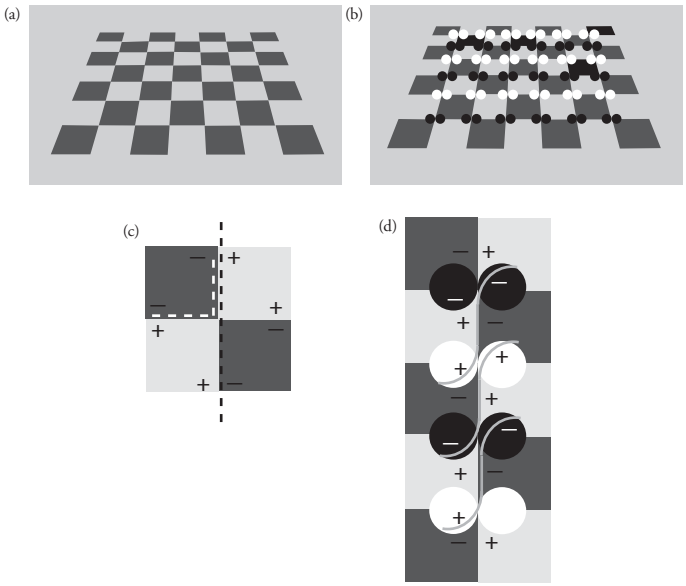


Figure II.32-4. (a) A checkerboard drawn in perspective. (b) Checkers are arranged on the checkerboard in (a). Although arranged at the opposite sides of lines converging at a vanishing point, the checkers appear to form alternating converging/diverging columns. (c) Directions of edges' completion (dotted lines). The line preserving the CP follows the square perimeter, while the line preserving the straight direction inverts the CP. (d) An enlargement of a portion of (b), in particular of the second, fourth, and sixth columns. The grey S-lines indicate the arches–edges conjunction path preserving the CP: a slight deviation from straightness obeys the Gestalt law of "good continuity." These paths also provide a local slant cue that in this case induces us to predict an illusory clockwise illusory tilt. The strength of this effect can be appreciated if we consider that the fourth column is inclined anticlockwise with respect to the vertical axis. The CP of the paths of arch–edge–arch conjunction alternate along the vertical axis. Moving in a downward direction, the polarities (left/right luminance variation) are the following: "+/–," "–/+," "+/–." The combination of the grey shades in the regions of surfaces convergence prevents the formation of symmetric paths of invariant CP. For example the "+/–" vertical edge preserves the CP if it joins to the top right black circle. A prolongation along the collinear edge or the top left black circle causes the inversion of the contrast sign. The same "+/–" vertical edge can preserve the CP in one direction only when prolonging at the lower extremity, that is, the one indicated in the top S-line. (Adapted from Parlangeli and Roncato (2010) with permission of Pion Publication.)

an object, moved to superimpose the meeting-point of converging lines, appears to be larger. In other words, layering the representation appears to be the solution by which two competing organizations in the perceptual field can coexist and continue to interfere with one another. A few unanswered questions remain concerning some aspects of this phenomenon.

THE CP RULE GENERALITY

In the version enunciated here, the CP rule cannot entirely account for some phenomena. Van Lier and Csathó (2006) demonstrated that contrast polarities across edges are not entirely sufficient to predict straightness misperceptions. According to their hypothesis, an explanatory framework must include the computation of contrast signs at a global level ("a polarity account dealing with certain global brightness differences would lead to better results"; p. 795). Parlangeli and Roncato (2010) demonstrated that the direction of luminance contrast cannot entirely account for some orientation misperceptions. And, in fact, in a balanced condition in which two paths of edge completion compete because both preserve their CP, the one coinciding with the border of higher contrast magnitude is nonetheless privileged. The CP rule must be integrated then by the following corollary: "where two alternative paths compete to emerge in opposite directions, the one with greater contrast luminance is likely to prevail."

THE BASIC DISTORTING PHENOMENON

It has been assumed that the local slant cues (i.e., the arch-edge-arch units along the vertical axes) are summed, leading to an overall tilt effect. The alternation in CP of these units (see the central axis of Fig. II.32-4d and its caption) leads us to note a close resemblance with the Fraser (1908) illusion effect. This illusion is perceivable when we draw a line of alternating black and white tilted segments in a grey background. The line appears to tilt in that direction. If the segments in alternating colors are replaced by isochromatic ones, the illusion weakens. The vertical axes in Figure II.32-4d can be assimilated to a Fraser line in which the segments are replaced by curved path of arch-edge-arch conjunction. The basic distorting effect underpinning the waving illusion of Figure II.32-1a likely coincides with the basic effect generating the Fraser illusion.

THE NATURE OF THE LOCAL SLANT CUES: A VIRTUAL COMPLETION?

We have used the term "path of conjunction" to designate a local slant cue, but it is not perceived as a figural unit. The contours of figural components (rectangles, disks) do not appear to reassemble into new figures following the paths indicated by dashed lines. We continue to perceive single units that are not linearly arranged as they are geometrically. A local slant is perceived at tile corners in Figure II.32-2a in the form of a deviation from straightness but nothing similar is found in Figures II.32-3 and II.32-4. If a common, basic effect is at play, this can only be an "invisible path" of conjunction of edges sharing the same CP. What was conjectured to account for the distorting effects in Figure II.32-4a—that is, a layered representation—may then be extended to other orientation misperceptions. In other words, inclined paths of edge conjunction complete and manifest their distorting effects in a perceptual layer that is different from the layer of checkerboard formation. The layer does not interfere in the contour formation but in the orientation perception.

THE CP RULE VIOLATIONS

A final remark must be made concerning the role of the CP rule. While it represents a "preferential" rule, it does not preclude the possibility that edges of opposite CPs can merge. Simple observation verifies the binding effect. Phenomenal transparency is obtained by combining opaque surfaces, and the figures that emerge are different from the component shapes. One of the more vivid transparency effects is obtained by a combining shades of grey that generate a contour in which CP is inverted. As Field, Hayes, and Hess (2000) have demonstrated, CP reversal does not prevent edges completion, although it renders the contour detectable to a lesser degree.

Alternative accounts to the one put forward here can be found in the studies of Morgan and Moulden (1986) and Kitaoka, Pinna, and Brelstaff (2004). Van Lier and Csathó (2006) and McCormick, Hon, Huang, and Altschuler (2012) have demonstrated some limits of CP rule application.

ACKNOWLEDGMENTS

Figure II.32-4 is a replica of some demonstrations that the author provided in an article published by *Perception*. The author is grateful to Pion Publications for the permission to reproduce.

REFERENCES

Benary, W. (1924). Beobachtungen zu einem Experiment über Helligkeitskontrast. *Psychologische Forshung*, *5*(1), 131–142.

Dresp, B. (1999). Dynamic characteristics of spatial mechanisms coding contour structures. *Spatial Vision*, *12*(2), 129–142.

Field, D. J., Hayes, A., & Hess, R. F. (2000). The role of polarity and symmetry in the perceptual grouping of contour fragments. *Spatial Vision*, *13*(1), 51–66.

Fraser, J. (1908). A new visual illusion of direction. *British Journal of Psychology*, *2*(3), 307–320.

Kitaoka, A., Pinna, B., & Brelstaff, G. (2004). Contrast polarities determine the direction of Café Wall tilts. *Perception*, *33*(1), 11–20.

Lier, R. van, & Csathó, Á. (2006). Dancing shapes: A comparison of luminance-induced distortions. *Perception*, *35*(6), 775–798.

McCormick, M. C., Hon, A. J., Huang, A. E, & Altschuler, E. L., (2012). Contrast polarity preservation's role in perception: Explained and unexplained stimuli, *Perception*, *41*(1), 12–25.

Morgan, M., & Moulden, B. (1986). The Münsterberg figure and twisted cords. *Vision Research*, *26*(11), 1793–1800.

Parlangeli, O., & Roncato, S. (2008). Boundary completion, contrast polarity, and the perception of illusory tilt. *Perception*, *37*(4), 535–556.

Parlangeli, O., & Roncato, S. (2010). Draughtsmen at work. *Perception*, *39*(2), 255–259.

Roncato, S. (2000). The effect of luminance variation on the apparent position of an edge. *Attention, Perception & Psychophysics*, *62*(4), 762–785.

Roncato, S., & Casco, C. (2003). The influence of contrast and spatial factors in the perceived shape of boundaries. *Attention, Perception & Psychophysics*, *65*(8), 1252–1272.

Chapter 33
Antigravity Slopes

A New Type of Visual Illusion

Kokichi Sugihara

SLOPE ILLUSION

Human vision sometimes perceives a slope to have an opposite orientation to the actual orientation. If we put a ball on such a slope, the ball appears to be rolling uphill as if it has defied the law of gravity. This is a type of visual illusion. We call this motion and slope an *antigravity motion* and *antigravity slope*, respectively.

Figure II.33-1 shows an example of an antigravity slope; (a) is an image taken at a specific viewpoint that generates an illusion, whereas (b) is an image of the same slope taken from a typical viewpoint. The orientation of the slope we perceive in (a) is opposite to that we perceive in (b). The actual orientation is the orientation we perceive in (b). Therefore, if we put a ball on the slope and observe it from the viewpoint in (a), we perceive antigravity motion; that is, the ball seems to be rolling uphill.

CONSTRUCTION OF AN ANTIGRAVITY SLOPE

The slope shown in Figure II.33-1 can be constructed with the aid of a computer. First, we generate a model of an ordinary slope consisting of a horizontal base plate, two vertical columns with different heights positioned on the base plate, and a slope fixed to the top of the two columns. This can be done using standard solid modeling techniques (Mortenson, 1997). As expected, the height of this slope decreases from the taller column to the shorter column.

Second, we fix a viewpoint and an image plane and project the slope onto the image plane to obtain an image of the slope. This can be done using a standard computer graphics technique (Foley, van Dam, Feiner, & Hughes, 1995). When we see this image, we perceive the orientation of the slope correctly; that is, the slope descends from the taller column to the shorter column.

Third, we reconstruct a three-dimensional solid structure from the two-dimensional image. However, the solid structure is not unique because the distance from the viewpoint to the object surface is lost in the image; infinitely many solid structures have projections that coincide with the image (Sugihara, 1986). Hence we have freedom in the choice of orientation of the slope and can choose a slope whose orientation is opposite to that of the original slope. An example of this slope is shown in Figure II.33-1. See Sugihara (2013) for the details of this procedure.

The base plate of the resulting solid structure is the same as the original one, but other parts are quite different: the columns are slanted instead of being vertical, and the top end of the originally taller column is lower than that of the other column, resulting in a slope that slants opposite to the original slope.

Employing the same method, we can construct many variants of antigravity slopes. Another example is the "Magnet-Like Slopes" shown in Figure II.33-2, where (a) shows the image taken at the viewpoint that generates an illusion and (b) shows the image taken at another viewpoint. This solid structure has four slopes supported by columns, and all of them look slanted from the highest center plate outward. If we place the balls on any of the four slopes, they appear to be rolling uphill; however, in reality the center is the lowest, as shown in Figure II.33-2b. This antigravity illusion won first prize at the Sixth Best Illusion of the Year Contest in 2010 (Sugihara, 2010).

EXPLANATION ON THE ANTIGRAVITY ILLUSION

The antigravity illusions shown in Figures II.33-1 and II.33-2 are the result of depth ambiguity. Images lack information about the depth, which is the distance from the viewpoint to the object surface, and hence our brains must guess the depth when interpreting the two-dimensional image as a three-dimensional solid structure. Mathematically, the possibilities are infinite, but our brain usually chooses only one depth and does not consider the other possibilities. This choice by the brain can be explained if we accept the following assumption:

> **Assumption 1.** The human brain prefers to interpret vertical columns in a two-dimensional image as being vertical in three-dimensional space to being slanted.

Once we believe that the columns are vertical, then it is a logical consequence that a longer column supports a higher part of the slope. This helps to explain why the antigravity illusion arises.

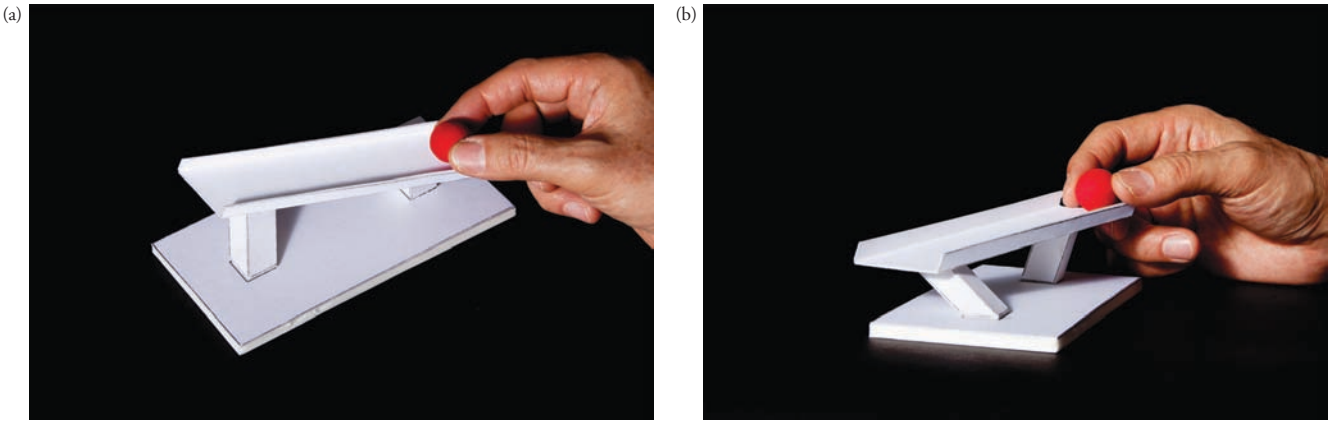


Figure II.33-1. Simple example of an antigravity slope. (a) Image of the slope seen from a specific viewpoint. (b) Image of the same slope seen from a typical viewpoint. (Kokichi Sugihara.)

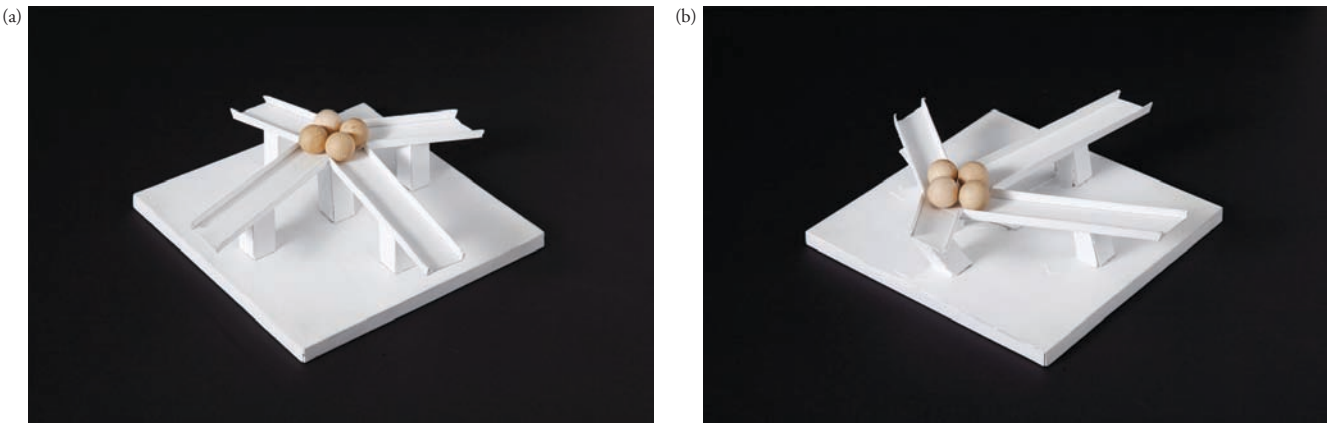


Figure II.33-2. "Magnet-Like Slopes," which won the first prize in the Sixth Best Illusion of the Year Contest, 2010. (a) Image of the solid structure seen from a specific viewpoint. (b) Image of the same solid structure seen from a typical viewpoint. (Kokichi Sugihara.)

ILLUSION OF A SLOPING ROAD

The slope illusion is also witnessed in real life. For example, sometimes a driver misunderstands the orientation of a road along which he or she is driving. Throughout the world, many sight-seeing spots have supposed "gravity hills" at which the orientations of the perceived roads are opposite the actual orientations. Two famous examples are Magnetic Hill in Moncton City, New Brunswick, Canada, and Spook Hill, in Lake Wales, Florida. Figure II.33-3 shows the Mysterious Road at Yashima Drive Way, Kagawa Prefecture, Japan; the nearer part of this road seems to be going downhill to most people, but it is actually going uphill (Bressan, Garlaschelli, & Barracano, 2013).

If a driver misunderstands the uphill road as downhill, he or she will not push the acceleration pedal and consequently the speed of the car will decrease. This is known as one of the sources of highway traffic congestion. Thus antigravity slopes, such as those in Figures II.33-1 and II.33-2, are designed in such a way that the lengths of the supporting columns mislead the viewer about the orientation of the slope. However, the road-slope illusion is not due to columns. Even if the real road was supported by columns, those columns would not be seen by drivers. Therefore, other reasons must explain the road-slope illusion.

Antigravity slopes without supporting columns can also be constructed in laboratories. Figure II.33-4 shows an example. Panel (a) gives an impression that the center is the highest, whereas panel (b) gives the opposite impression. However, the truth is that panels (a) and (b) are the same solid structure seen from two different angles. This solid structure is obtained by removing the supporting



Figure II.33-3. Yashima Drive Way in Shikoku Island, Japan. The nearer part of this road is going uphill although it looks downhill to most people. (Kokichi Sugihara.)

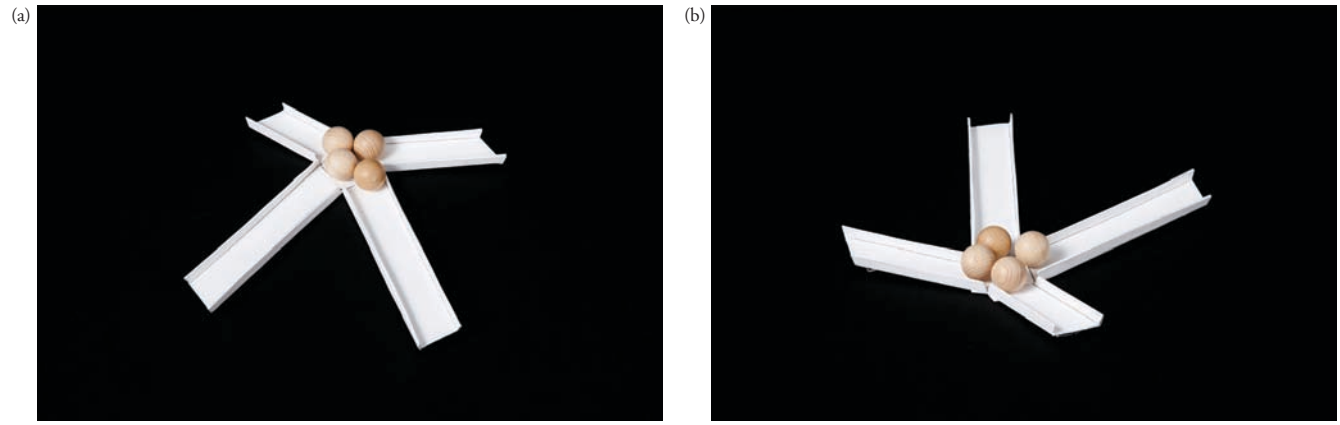


Figure II.33-4. Antigravity slopes without supporting columns; (a) and (b) are the same solid obtained from "Magnet-Like Slopes" in Figure II.33-1 by removing the supporting columns. (Kokichi Sugihara.)

columns from the "Magnet-Like Slopes" shown in Figure II.33-2. The central part is a little lower, and all four slopes go downhill gently toward the center. From this example we can assert that antigravity slopes can be constructed even when no supporting columns are used.

The next question is then: "Why does an illusion arise from this solid structure?" This question can be answered if we accept the next assumption.

> **Assumption 2.** The human brain prefers the most symmetric shape as the interpretation of a two-dimensional image.

The solid structure in Figure II.33-4 can be interpreted as a symmetric shape in the sense that four roads cross at right angles at the center; that is, the shape is symmetric with respect to the rotations of 90° around the vertical axis. This interpretation logically leads us to conclude that the center is the highest in (a), while it is the lowest in (b). Thus we can explain why (a) and (b) in Figure II.33-4 generate impressions of mutually different shapes.

The Mysterious Road in Figure II.33-3 can also be explained by Assumption 2. The road in the figure consists of two parts (a nearer part and a farther part) with different slopes. One symmetric interpretation is that this shape is mirror-symmetric with respect to the vertical plane. This interpretation logically leads one to conclude that the nearer part is downhill and the farther part is uphill. In this way, the illusion of antigravity slopes can be explained by Assumptions 1 and 2.

REFERENCES

Bressan, P., Garlaschelli, L., & Barracano, M. (2013). Antigravity hills are visual illusions. *Psychological Science, 14*(5), 441–449.

Foley, J. D., van Dam, A., Feiner, S. K., & Hughes, J. F. (1995). *Computer graphics: Principles and practice* (2d ed.). Reading, MA: Addison-Wesley.

Mortenson, M. E. (1997). *Geometric modeling* (2d ed.). New York: Wiley.

Sugihara, K. (1986). *Machine interpretation of line drawings*. Cambridge, MA: MIT Press.

Sugihara, K. (2010). Impossible motion: Magnet-like slopes. Winner: Best illusion of the year contest. http://illusionoftheyear.com/2010/impossible-motion-magnet-like-slopes/

Sugihara, K. (2013). Design of solids for antigravity motion illusion, *Computational Geometry: Theory and Applications*, *47*(6), 675–682.

Chapter 34
The Geometric-Optical Illusions of J. J. Oppel

Dejan Todorović

INTRODUCTION

The term "geometric-optical illusions" probably first appeared in print in the title of a series of three papers by Johann Joseph Oppel (1815–1894), published in the journal *Jahresbericht des Physikalischen Vereins zu Frankfurt am Main* (Annual Report of the Frankfurt/Main Physical Society). Although Oppel is often credited as the originator of the academic study of visual illusions, the actual content of these papers (Oppel, 1854–1855, 1856–1857, 1860–1861) is not well known. Rare recent mentions of his work include Westheimer (2008), Vicario (2011), and Phillips and Wade (2014), who also provided interesting biographic and scientific details. Oppel's name is usually associated only with the Oppel–Kundt illusion, but this effect was only briefly mentioned in his third paper, without any illustrations and not in the form in which it was later studied (see chapter 35 in this volume). However, in his papers he discussed a variety of other effects in some detail and presented a total of 29 figures. In this chapter I argue that some of these effects are important but neglected precursors of certain classical illusions generally not associated with his name. Variations of Oppel's designs are also presented, in which I have attempted to strengthen the illusory effects or make them more salient.

The three illusion papers comprise only a small portion of Oppel's versatile scientific output (Phillips & Wade, 2014). More than 20 references to his papers, dealing mainly with color, motion, and stereo vision, are listed in the bibliography included in the second edition of Helmholtz's *Treatise on Physiological Optics* (Helmholtz, 1896). He also published papers in acoustics, devised apparatuses for educational demonstrations of physical phenomena and color blindness, and wrote a mathematics textbook. An example of his wide-ranging interests is a short paper titled "The Acoustical Aspects of the Call of the Cuckoo," in which he reported that in his nature trips he judged the absolute pitches and musical intervals of cuckoo calls (descending two-tone sequences) on no less than 157 occasions (Oppel, 1871). He also contributed a large amount of material on the Frankfurt dialect of German, for the collection of which he devised a special shorthand script.

THE GEOMETRICAL-OPTICAL ILLUSIONS

In his introductory remarks Oppel states that he chose the name "geometrical-optical illusions" to denote a class of phenomena in which the eye or, more appropriately, the unconscious judgment of the reason (he explicitly used this phrase, which is usually associated with Helmholtz) is deceived with respect to relations of dimensions and directions, that is, magnitudes of lines and angles. These phenomena, he says, are clearly different from what are usually called optical illusions, such as in the salient apparent size of the rising moon. These usual optical illusions, according to Oppel, are characterized by the relations of three quantities—apparent size (visual angle), real size, and distance of objects—and they are based on the fact that if one of a pair of these magnitudes is wrongly set too high or too low, the third one is judged incorrectly. In contrast, in the class of geometrical-optical illusions, neither the distance nor the visual angle has any effect.

Oppel begins his account with what he says is an easily made observation, especially for teachers of geometry: that vertical dimensions generally appear larger than corresponding equal horizontal dimensions. His examples do not include the now well-known inverted-T figure but rather a variety of other cases. Thus he writes that a rectangle with height of 8 inches and base of 8.5 inches looks like a square to many casual viewers, whereas a nearby actual square looks too tall by about half an inch. Similarly, a horizontal ellipse whose ratio of axes is 17:18, or even 11:12, looks like a circle but is seen as an ellipse once it is turned by 90°. An angle of 93° looks like a right angle if its orientation is such that its imaginary bisector line is vertical. However, if the angle is rotated by 90°, so that its bisector is horizontal and it is open toward the left or right, it is immediately seen as obtuse. Furthermore, an angle of 87° in the last mentioned orientation looks like a right angle but is immediately seen as acute if turned by 90°. When asked to draw an isosceles right triangle with horizontal base, most students draw an obtuse angle at the apex, whereas if the base is vertical they draw an acute angle. Furthermore, oblique lines look longer than equally long horizontal lines. Namely, when asked to draw an equilateral triangle with horizontal base, students tend to draw the base line too long; if the blackboard is

rotated such that one of the two other sides of the triangle is horizontal, the base is immediately seen as too long.

A final example involves the judgment of the diameter of a cylinder, such as a drinking glass, in the following way, akin to a parlor trick. One is asked to observe the glass when it is oriented vertically, and the task is to judge how tall it will be when it is turned over into the horizontal position. The judgment is executed by forming a corresponding vertical span between one's thumb and index finger. According to Oppel, when the glass is turned, one finds that the manual judgment falls short of the actual diameter. Oppel claimed that the explanation of all these phenomena is "without doubt" the fact that the eye is accustomed to see much larger extents in the horizontal than in the vertical orientation and thus out of habit slightly overestimates the vertical extents. In my judgment the explanation is not particularly compelling, because much the same account could have been invoked if horizontal rather than vertical extents had been found to be overestimated.

In the following I present only a part of Oppel's wide-ranging observations collected in the three papers. They include three sets of cases that can plausibly be regarded as versions of illusions that were later published by other authors.

PRECURSORS OF THE MÜLLER-LYER ILLUSION

Oppel was a high-school teacher of mathematics and physics. Teaching these subjects provided quite a few occasions for him and his colleagues in other Frankfurt high schools to note discrepancies between physical and perceived features of various geometrical constructions. An example is Figure II.34-1a, a diagram illustrating a geometrical procedure to divide the circumference of a circle into

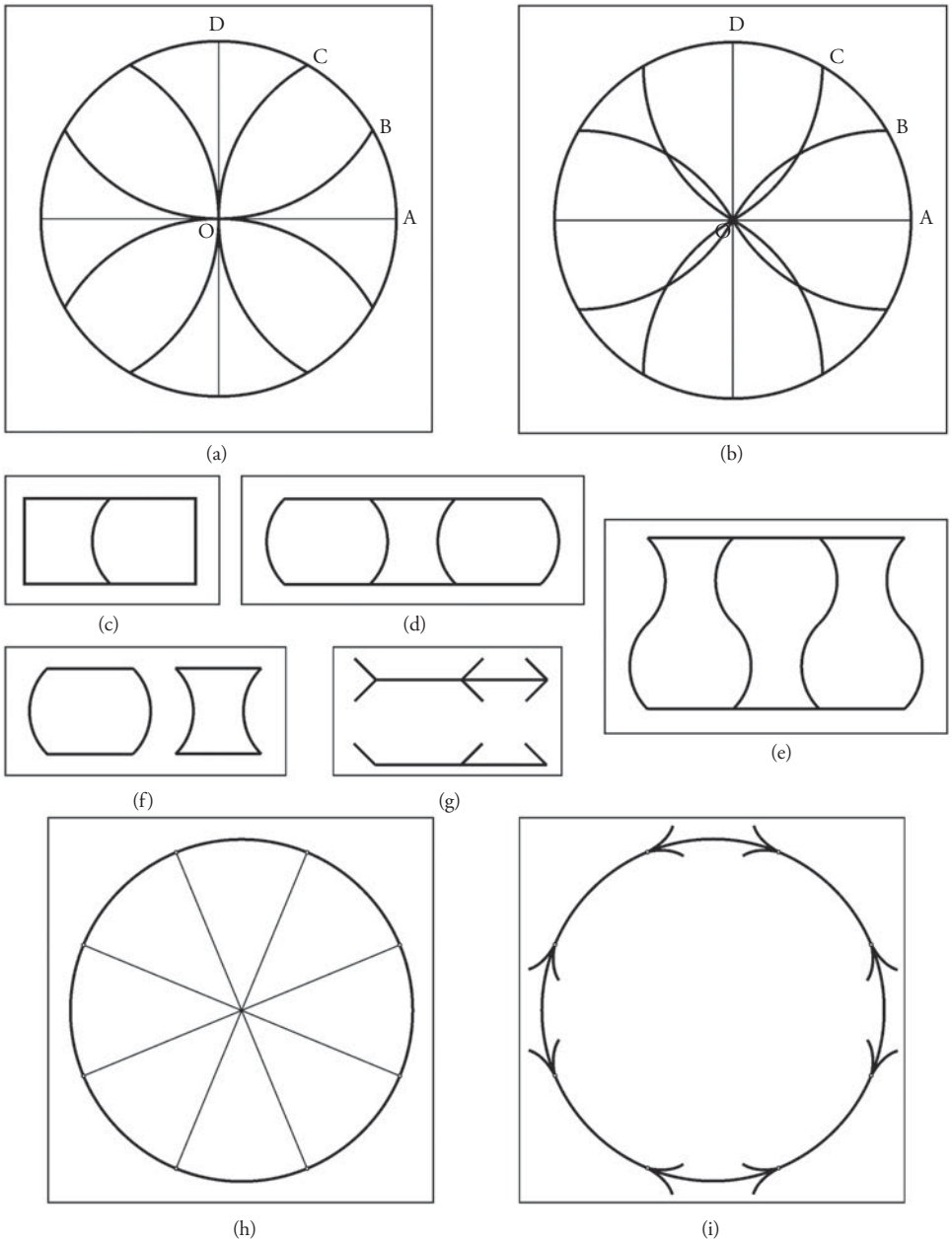


Figure II.34-1. Oppel's designs as precursors of the Müller-Lyer illusion. (a) through (e): After Oppel (1854–1855), Fig. 11, Fig. 12, Fig. 13, Fig. 14, and Fig. 15, respectively. (f) after Wundt (1898), Fig. 53 (g) based on Thiéry (1896), Fig. 45, (h), (i) Author's representations. Credit: Author's representations.

12 equal 30° segments (such as AB, BC, CD, etc.) by drawing, within that circle, four circular arcs with the same radius as the circle, with centers at its four poles. Oppel relates that he often presented this figure in his geometry lectures but that the students "did not feel any incentive to try to prove the correctness of this construction" because the segments did not actually look equally long; in particular, the middle segment in each of the four quadrants (such as BC in the first quadrant) looked longer than the two flanking segments (AB and CD). Oppel's explanation of this effect was that the eye confounds the lengths of the segments with the areas of the regions of which they form the borders; indeed, the area of the region corresponding to the middle segment (OBC) is larger than the areas of the regions corresponding to the flanking segments (OAB, OCD). For corroborating evidence serving as control, he presented Figure II.34-1b as a "counter-trial" (*Gegenversuch*), in which the same arcs were used but with opposite curvature, with the consequence that the illusion disappeared or inverted. As additional evidence he presented much simpler displays, such as Figure II.34-1c, in which an arc connects the midpoints of the top and bottom side of a rectangle but the left halves of the sides look shorter than the right halves. According to Oppel, this is due to the fact that the left region has a smaller area than the right region. A similar explanation was offered for Figure II.34-1d. Here the top and bottom sides of the rectangle are divided into three equal segments joined by two arcs, but the middle segments look shorter than the flanking segments. As another demonstration Oppel presented the elegant Figure II.34-1e, in which, like in Figure II.34-1d, the top and bottom horizontal lines are divided into three objectively equal segments that look different. However, he appears to have overlooked the fact that his explanation does not apply for this display, because all three vase-shaped regions have the same area.

Oppel's designs bear a certain relation to the famous configuration introduced by F. C. Müller-Lyer 35 years after Oppel's paper. This relation was noted by Wilhelm Wundt, who presented the two designs in Figure II.34-1f, which are components of Figure II.34-1d, to argue that that they involve Müller-Lyer "motives." Indeed, in one of the designs the top and bottom corners are equivalent to portions of Müller-Lyer patterns with outer chevrons, whereas in the other design they are equivalent to portions of Müller-Lyer patterns with inner chevrons. There is a difference from the standard pattern in that the chevrons are curvilinear, but such Müller-Lyer patterns were already found to be effective by Brentano (1892). Armand Thiéry (1896), a student of Wundt, studied experimentally whole and partial (one-sided) Müller-Lyer configurations as in Figure II.34-1g and found that the illusory effect in the partial configuration was in the same direction as in the whole configuration but weaker. Later investigations confirmed this result (Predebon, 1996; Warren & Bashford, 1977). It thus seems fair to conclude that Oppel's designs were precursors to the Müller-Lyer figures. Müller-Lyer himself was apparently not aware of Oppel's work, since he did not refer to Oppel in his papers (Müller-Lyer, 1889, 1896a, 1896b).

The illusory effect in Figure II.34-1a can be strengthened and made more salient by some reorganization, which includes completing the one-sided Müller-Lyer "motives" and removing extraneous elements. Figure II.34-1h presents a division of the area and circumference of a circle into eight equal segments. In Figure II.34-1i curved chevrons were added to this figure at eight locations on the circumference indicated by the small circles (see also https://visiome.neuroinf.jp/modules/xoonips/detail.php?item_id=6880). In this streamlined version of Figure II.34-1a the illusory difference in length between the segments flanked by inner and by outer chevrons is rather apparent. It is not clear how perspective-based accounts of the Müller-Lyer illusion could explain the illusory effect in this pattern.

PRECURSORS OF POSITION ILLUSIONS OF THIÉRY, WUNDT, AND JUDD

Another illusion inspired by a geometry exercise for high-school students is illustrated in Figure II.34-2a. The task was to draw the two diagonals of a square and then mark, not by ruler but by visual judgment, the midpoints of each of the four half-diagonals; the correct positions of the marks are indicated in the figure. A colleague of Oppel noted, and Oppel confirmed, that markings by students tended to deviate from the actual midpoints toward the direction of the center of the square. From this result it follows that the locations of the marks in Figure II.34-2a, which are positioned at the actual midpoints, will appear shifted toward the corners of the square. On the other hand, when the task was to mark the midpoints of the extents between the center of the square and each of its four sides, as in Figure II.34-2b, the marks tended to be placed nearer to the sides of the square than to its center. Consequently, the marks in Figure II.34-2b, which are positioned at the actual midpoints, appear nearer to the center than to the sides. However, no illusory shift was found when the same task as in Figure II.34-2b was presented for a square with lines connecting the midpoints of opposite sides, as in Figure II.34-2c. Note that in both Figures II.34-2a and II.34-2b the perceived positions of the halfway marks tend to be shifted toward the corner of a triangle, whereas in Figure II.34-2c no triangle is present. Oppel noted, without presenting an illustration, that a simplified version of this effect involved a single equilateral triangle with a horizontal base; given the task to place a point at the middle of its height, subjects would set it too low.

Figure II.34-2d presents a version of this illusion introduced 40 years after Oppel by Thiéry (1896), who did not cite Oppel's work. In this figure the horizontal line passes through the midpoint of the height of the triangle but appears higher than halfway. In order to make it appear at half height, Thiéry's subjects had to lower its actual height; this illusory effect increased as the size of the top angle of the triangle decreased. Thiéry remarked that this fact was familiar to typesetters, who knew that in the capital letter A the horizontal line had to be placed lower in order to appear at half height. From Thiéry's results it follows that the effect in Oppel's designs would be stronger if the angles of the triangles were smaller. The effect indeed seems to be somewhat stronger in my designs in Figures II.34-2e and II.34-2f, which are counterparts of Figures II.34-2a and II.34-2b but contain triangles with smaller angles at the apexes.

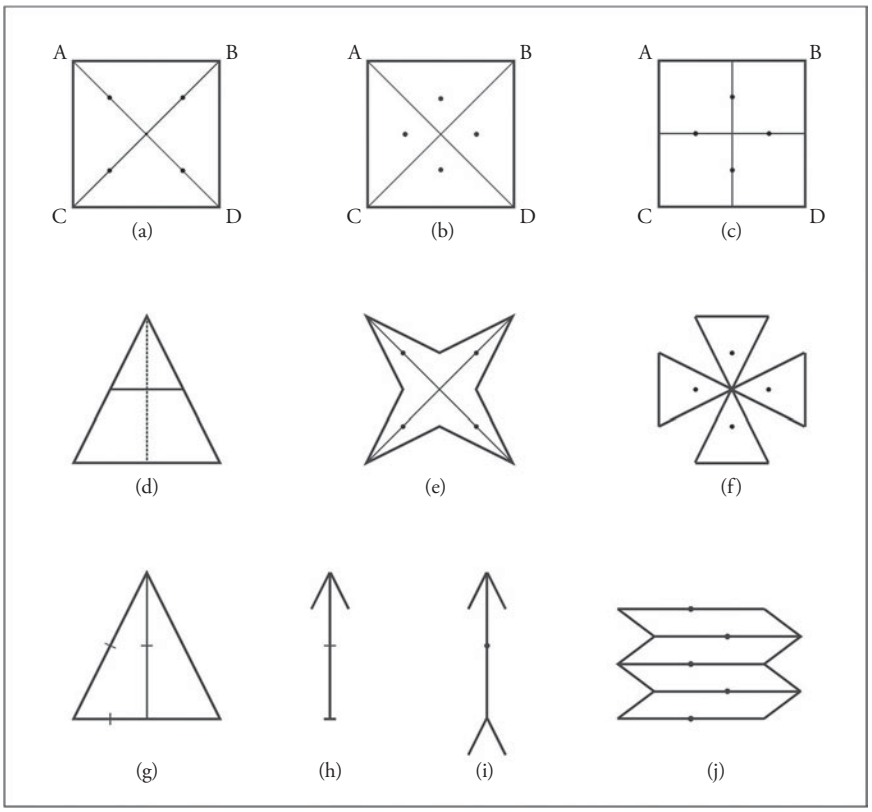


Figure II.34-2. Oppel's designs as precursors of the illusions by Thiéry (1896), Wundt (1898), and Judd (1899). (a) through (c): After Oppel (1856–1857), Fig. 3, Fig. 4, and Fig. 5, respectively. (d) through (f): Author's representations. (g) and (h): After Wundt (1898), Fig. 54 and Fig. 55, respectively. (i) After Judd (1899) (j) Author's representation. Credit: Author's representation.

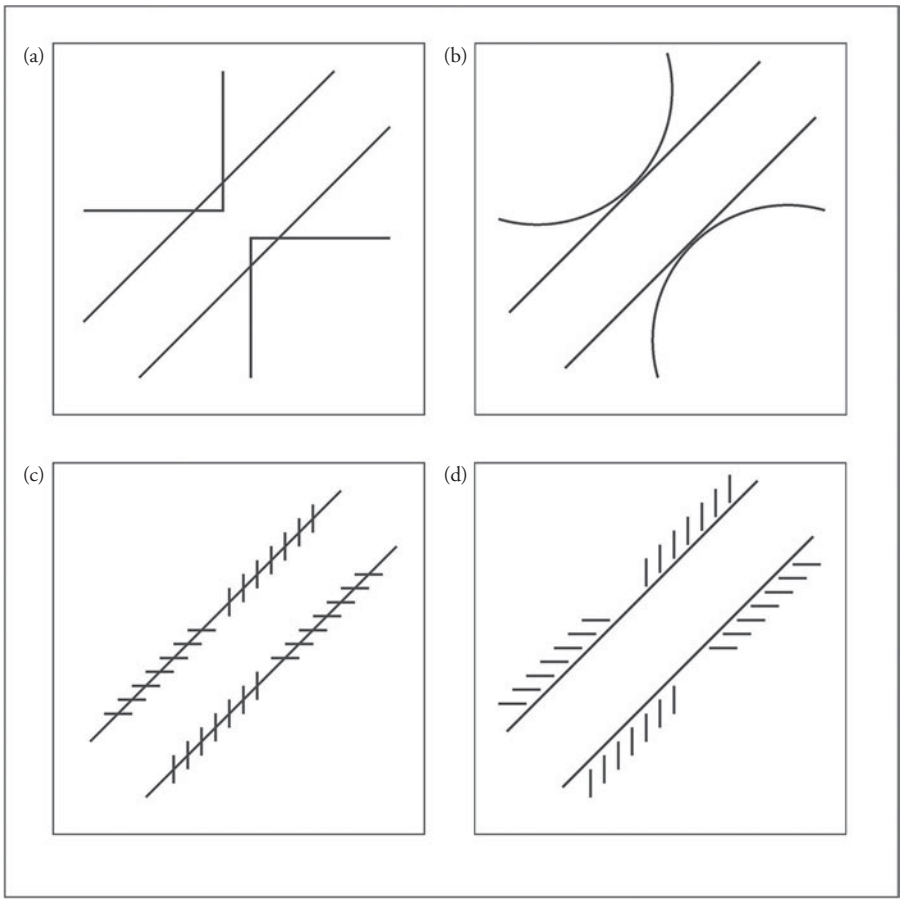


Figure II.34-3. Oppel's designs as precursors of the Zöllner illusion. (a) and (b): After Oppel (1854–1855), Fig. 3 and Fig. 5, respectively. (c) and (d): Author's representations. Credit: Author's representations.

Thiéry's mentor Wundt (1898) explicitly credited Oppel (1856–1857) for this phenomenon and suggested that it was related to the Müller-Lyer illusion. Wundt contributed Figure II.34-2g, in which midpoints are indicated for the height of the triangle, for one of its legs, and for the half-base. He noted that in all cases the midpoint was perceptually shifted toward the nearest acute angle (or, in the case of the leg, toward the more acute of the two equally distant angles) but that the effect was largest for the height. He also contributed Figure II.34-2h, whose purpose was to show that to generate the illusory position shift, there was no need for the whole triangle to be present; an angle suffices. In recent years this illusion has been studied by Anstis, Gregory, and Heard (2009) and Day and Kim (2010).

Figure II.34-2h is very similar to Figure II.34-2i, which was presented by Judd (1899); the perceptual mislocalization of the midpoint is known as the Judd illusion (Köhler, 1947; Morgan, 1969; van der Kamp, Withagen, & de Wit, 2013). A "multiplied" Judd illusion is presented in Figure II.34-2j. All dots are located at the midpoints of the lines, but they look shifted rightward or leftward, according to the directions of the chevrons. It is plausible that all these illusions are mutually related, and also related to the Müller-Lyer illusion, but these relations need to be studied more.

PRECURSORS OF THE ZÖLLNER ILLUSION

A colleague of Oppel noted that when two lines forming an obtuse angle cross a straight line, the latter looks as if slightly bent or broken. This is illustrated in Figure II.34-3a (he recommended an obtuse angle of at least 135°, but I find that a 90° angle suffices). Note that the distance between the two geometrically parallel lines seems slightly wider in the region where they are crossed by the right angles, so that they appear to bulge out like a barrel and do not look like parallel straight lines. A similar small bulging effect appears when the two parallel lines are almost touched by circular arcs, as in Figure II.34-3b. The purpose of my Figure II.34-3c, in which the right angles in Figure II.34-3a are replaced by sets of short straight lines that cross the parallel lines at multiple places instead of just two, is to suggest that the illusory effect in Figure II.34-3a is related to the Zöllner illusion. Figure II.34-3d involves a more speculative suggestion that the effect in Figure II.34-3b may also be related to the Zöllner illusion, since the illusion-inducing effect of the short lines may be present even when they do not intersect the parallel lines (Earle & Maskell, 1995).

ACKNOWLEDGMENTS

This research was supported by the Serbian Ministry of Science Grant ON179033.

REFERENCES

Anstis, S., Gregory, R., & Heard, P. (2009). The triangle-bisection illusion. *Perception*, *38*, 321–332.

Brentano, F. (1892). Über ein optisches Paradoxon. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, *3,* 349–358.

Day, R. H., & Kimm, A. C. (2010). Analysis and explanation of the Thiéry–Wundt illusion. *Perception*, *39*, 942–952.

Earle, D. C., & Maskell, S. J. (1995). Spatial filtering and the Zöllner–Judd geometrical illusion: Further studies. *Perception*, *24*, 1397–1406.

Helmholtz, H. (1896). *Handbuch der physiologischen Optik* (2nd edition). Hamburg: Voss.

Judd, C. H. (1899). A study of geometrical illusions. *Psychological Review*, *6,* 241–261.

Köhler, W. (1947). *Gestalt psychology.* New York: Liveright.

Morgan, M. J. (1969). Estimates of length in a modified Müller-Lyer figure. *The American Journal of Psychology*, *82*, 380–384.

Müller-Lyer, F. C. (1889). Optische Urteilstäuschungen. *Archiv für Anatomie und Physiologie, Physiologische Abteilung 2*(Suppl.), 263–270.

Müller-Lyer F. C. (1896a). Über Kontrast und Konfluxion (zweiter Artikel). *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, *9*, 421–431.

Müller-Lyer F. C. (1896b). Zur Lehre von den optischen Täuschungen: Über Kontrast und Konfluxion. *Zeitschrift für Psychologie*, *9*, 1–16.

Oppel, J. J. (1854–1855). Über geometrisch-optische Täuschungen. *Jahresbericht des physikalischen Vereins zu Frankfurt am Main,* 37–47.

Oppel, J. J. (1856–1857). Nachlese zu den "geometrisch-optischen Täuschungen." *Jahresbericht des physikalischen Vereins zu Frankfurt am Main,* 47–55.

Oppel, J. J. (1860–1861). Über geometrisch-optische Täuschungen (Zweite Nachlese). *Jahresbericht des physikalischen Vereins zu Frankfurt am Main*, 26–37.

Oppel, J. J. (1871). Der Kukuksruf in akustischer Beziehung. *Annalen der* Physik, *220*(10), 307–309.

Phillips, D., & Wade, N. J. (2014). The elusive Johann Joseph Oppel (1815–1894). *Perception*, *43,* 869–872.

Predebon, J. (1996). The role of the angle components in the wings-in and wings-out forms of the Müller-Lyer illusion. *Perception*, *25*, 773–782.

Thiéry, A. (1896). Ueber geometrisch-optische Täuschungen, Schluss. *Philosophische Studien*, *12,* 67–126.

van der Kamp, J., Withagen, R., & de Wit, M. M. (2013). Cultural and learning differences in the Judd illusion. *Attention, Perception, & Psychophysics*, *75*(5), 1027–1038.

Vicario, G. B. (2011). *Illusioni ottico-geometriche: una rassegna di problemi.* Venezia, Italy: Istituto Veneto di Scienze, Lettere ed Arti.

Warren, R. M., & Bashford, J. A. (1977). Müller-Lyer illusions: Their origin in processes facilitating object recognition. *Perception*, *6*, 615–626.

Westheimer, G. (2008). Illusions in the spatial sense of the eye: Geometrical-optical illusions and the neural representation of space. *Vision Research*, *48*(20), 2128–2142.

Wundt, W. M. (1898) Die geometrisch-optischen Täuschungen. *Ahandlungen der königlich sächsischen Gesellschaft der Wissenschaften, mathematisch-physische Klasse*, *42*(2), 55–178.

Chapter 35
The Oppel–Kundt Illusion

Jiří Wackermann

INTRODUCTION

The Oppel–Kundt illusion (OKI) belongs to the class of so-called geometric-optical illusions: a part of the visual field, subdivided by an array of visual elements, appears larger than undivided space of the same extent. This can be illustrated by simple visual displays such as Figure II.35-1. There is a great variety of graphic representations of the OKI, using different visual elements and also differing in the magnitude of the illusory effect. The OKI is also known as "illusion of interrupted extent" (Luckiesh, 1922; Sanford, 1903) or "filled space illusion" (Lewis, 1912).

The illusory effect can be easily measured using the reproduction method: a movable element delimiting the "empty" part of the figure is adjusted by the observer to match the distance between the delimiters of the "filled" part (Fig. II.35-2). (The two-alternatives forced-choice method can be applied in order to determine the point of subjective equality of the two extents: a more laborious and less frequently used procedure.) As the undivided and the subdivided extents are subjectively equal, the illusory effect is quantified by the excess of the adjusted extent relative to the geometrically correct response. The average effect magnitudes reach typically up to 15% of the reference extent (Wackermann & Kastner, 2009, 2010) and may be, in individual cases, even higher. Bulatov, Bertulis, and Mickienė (1997) demonstrated a linear dependence of the illusory overestimation on the extent of the reference part in the visual field, a fact that advocates the use of relative expansion to measure the effect. Yet many authors report effect magnitudes in absolute length or angular units for their particular stimulus configuration rather than in relative numbers, which makes comparisons between different studies difficult.

HISTORY

The discovery of the illusion is attributed to the German physicist J. J. Oppel, who coined the notion of geometric-optical illusions (Oppel, 1854–1855). In fact, the phenomenon is mentioned in Oppel's (1860–1861) second communication on geometric-optical illusions, where he reported an asymmetry of freehand drawings of rectangles made by his subjects on sheets of lined paper. In these experiments, the observed expansion effect was entangled with another illusion, that of horizontal versus vertical extent. Oppel (1860–1861) summarized his observation as follows: "a multiple division of an area or of a [linear] dimension makes the latter to appear slightly larger than an undivided area or line of the same extent" (p. 35), but he did not author any of the figures circulating in the literature. Independently from Oppel, Hering (1861) demonstrated the illusory expansion in a series of dot displays similar to Figure II.35-1a (along with other visual phenomena, such as the Hering illusion). Kundt (1863) carried out measurements of the effect, using the reproduction method, and proposed an explanatory theory similar to that of Hering (see the Explanations and Models section). Aubert (1865) criticized Kundt's theory and pointed out discrepancies between Kundt's and his own measurements. Helmholtz (1867) showed Figure II.35-1b, referring to Kundt's experiments but not mentioning Oppel's report.

The naming of the phenomenon (Boring, 1942; Coren & Girgus, 1978; Robinson, 1998) is thus historically not quite correct: Oppel described a phenomenon similar but not identical to the form of the OKI known later, and Kundt did not discover the illusion but was probably the first to study it quantitatively. The reduction of the stimulus figure to a one-dimensional array of visual elements, partly used for illustrative purposes and partly enforced by experimental techniques, is due to several independent authors.

Except for a short episode in the early nativism versus empiricism debate (cf. Turner, 1997; see also the Explanations and Models section), the OKI has not played any important role in discussions or theories of visual illusions. The OKI is a pure "illusion of extent," containing no angular elements—such as the "wings" in the Müller-Lyer figure or the lateral tilted lines in the Ponzo figure—and is thus immune to explanatory approaches based on alleged distance cues (Gregory, 1963). Compared to other geometric-optical illusions, the OKI attracted relatively little attention of researchers and remained for decades underrepresented in the literature. There was, however, increasing interest in the OKI in the 1990s in studies modeling visual perception in terms of information processing or underlying neurophysiology (Bulatov et al., 1997; Erdfelder & Faul, 1994). Recently, the OKI has received even more attention in connection with the visual space anisometry observed in neurological patients (Doricchi et al.,

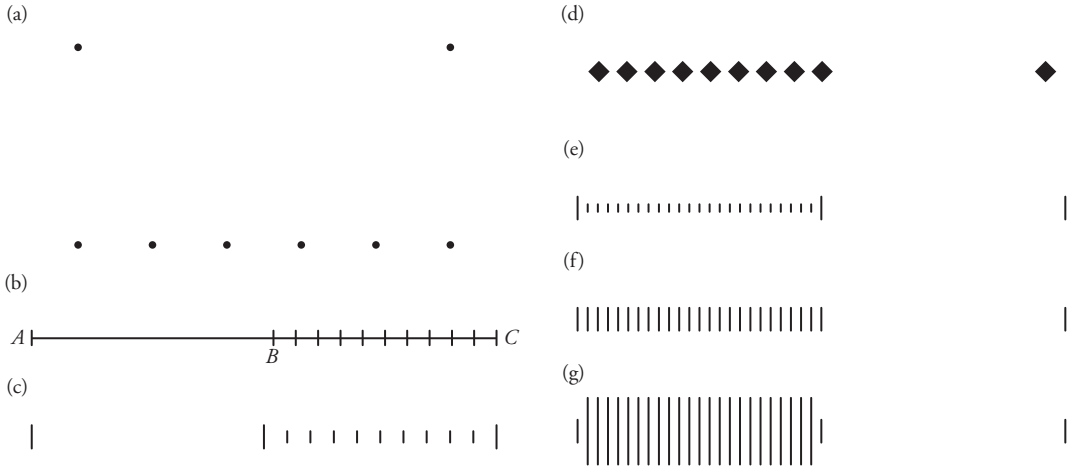


Figure II.35-1. Illustrations of the Oppel–Kundt illusion. (a) According to Hering (1861): the distance between the two dots in the upper row seems to be smaller than the distance between the two respective dots in the lower row. (b) According to Helmholtz (1867, Fig. 172): the distance *AB* seems to be equal to the distance *BC*, although objectively *AB* > *BC*. (c) The standard form shown in most reference works: the distance between the central and the left terminating stroke seems to be smaller than the distance between the central and the right terminating stroke, although objectively the two distances are equal. (d) The same illusion is obtained with a different appearance of the delimiters/subdividing elements (Wackermann, 2010). (e, f, g) Various forms of the stimulus figure demonstrating the paradoxical form of the OKI (Wackermann & Kastner, 2009).

2008; Ricci, Pia, & Gindri, 2004; Savazzi, Frigo, & Minuto, 2004; Savazzi, Posteraro, Veronesi, & Mancini, 2007). OKI-like arrays of linear elements may be used to induce such metric distortions experimentally, and, inversely, insights into principles determining the metric of the visual field may help to understand causal mechanisms involved in such neurological disorders.

DETERMINING FACTORS

The multiplicity of names under which the OKI is known (see Introduction section) indicates a variety of assumptions concerning the putative "causes" of the effect. None of these, however, reflects the complex interplay of determining factors revealed by further experimental research.

Early studies focused on the number of elements n subdividing a given spatial extent as the major determinant of the phenomenon. Most studies found an increasing effect magnitude for n increasing up to a certain value from where the effect magnitude slowly decreases (Bulatov et al., 1997; Knox & Watanabe, 1894; Lewis, 1912; Obonai, 1933; Piaget & Osterrieth, 1953; Spiegel, 1937; Wackermann & Kastner, 2010). Note that results for $n = 1$ (i.e., halving the subdivided part) are controversial (cf. Coren & Girgus, 1978; Spiegel, 1937). This nonmonotonicity of the illusory effect with regard to n is the most robust finding on the OKI. There is, however, no constant n_{max} for which the maximum effect occurs. The form of the response curve (i.e., the effect magnitude expressed as a function of n) depends on many procedural parameters, including optical properties of the stimulus, geometrical form of the subdividing elements, and so on.

A possible role of the appearance of dividers was suggested early (Messer, 1876) but systematically studied only much later. Spiegel (1937) used a horizontal array of short vertical lines, with delimiting and subdividing elements of the same height, as a standard version of the OKI figure. Shortening the vertical extent of the dividers reduced the effect magnitude, while their widening had no effect. Wackermann and Kastner (2009) observed that after extending the dividers further the effect magnitude did not increase or saturate but dropped to zero (the "paradoxical form" of the OKI). A follow-up study showed that, for fixed n, the effect magnitude is a continuous, nonmonotonic function of the dividers' height (Wackermann & Kastner, 2010). The OKI is thus not exclusively an effect of the subdivision but is multifactorially determined (Figures II.35-3a and II.35-3b): a finding that establishes connection between the OKI and "filled space" effects observed in two-dimensional configurations (see Related Phenomena section). However, the OKI is not a pure effect of filling either; this is demonstrated

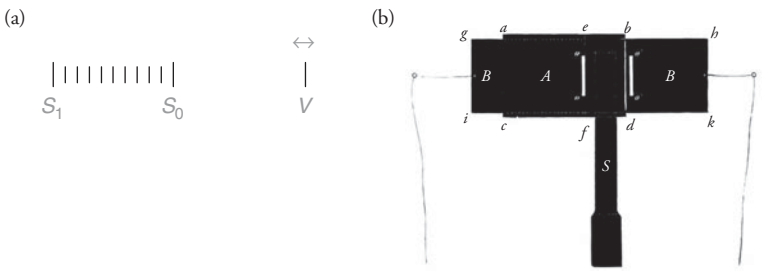


Figure II.35-2. (a) Principle of the reproduction method; the relative illusory effect is defined as $r = (v - s) / s$, where $v = VS_0$, $s = S_0S_1$ (Wackermann & Kastner, 2009). (b) Apparatus used by Aubert (1865, Fig. 48) to measure the illusory effect in an array of vertical strokes.

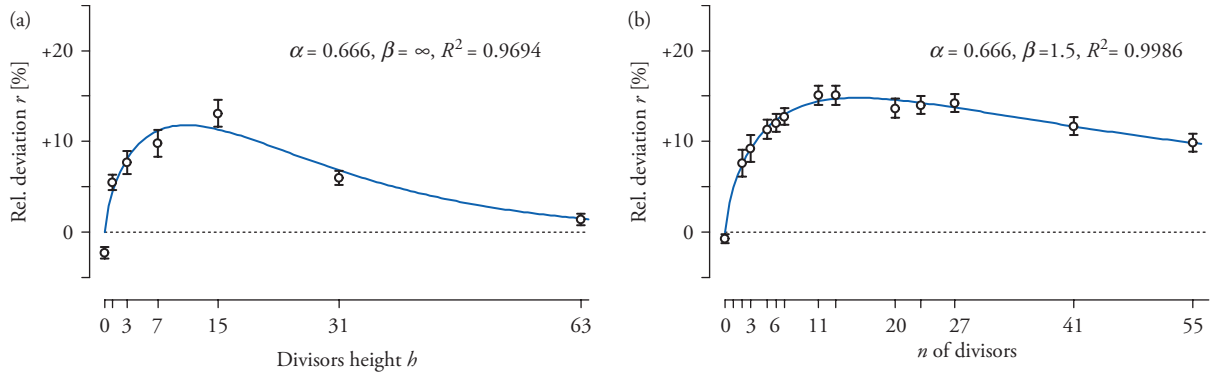


Figure II.35-3. (a) Effect magnitude as a function of the number of dividers n with fixed dividers' height h. (b) Effect magnitude as a function of dividers' height h with fixed number of dividers n. α and β are parameters of the data fitting functions (blue curves); R^2 is a measure of the goodness of fit. (From Wackermann & Kastner, 2010, with permission from the Nencki Institute of Neurobiology, Warsaw, Poland.)

by experiments in which the form and/or number of the subdividing elements varied while the relative coverage of the spatial extent with the "optical matter" was kept constant (Wackermann, 2012a).

As for the dependence of the OKI effect on other properties of the stimulus, data are rather scarce. Spiegel (1937) used luminous figures on a dark background as the standard condition; a reversal to a black figure on a white background reduced the effect magnitude slightly, while the overall form of the response curve remained unchanged. Decreased brightness of the luminous dividers, or a change of color from white to red, resulted in a minor decrease of the effect. Dworkin and Bross (1998) also reported increasing effect magnitudes with increasing brightness and higher contrast of the stimuli. Wackermann (2012b) found systematic effects of both contrast and visual polarity of the stimulus, contributing up to about 30% of the total effect magnitude. These results are important as they reveal the modulating role of optical properties of the stimulus and thus involvement of early-stage sensory processes in the illusion. Exposure times in tachistoscopic presentations of the stimulus figure seem to affect the illusion magnitude, but there is no perfect agreement between results (cf. Dworkin & Bross, 1998; Piaget & Bang, 1961; Surkys & Bertulis, 2010). Hemispheric functional specialization and interhemispheric differences in processing times may also play a role (Rothwell & Zaidel, 1990).

Most of the results reported here were obtained with the "standard" orientation of the stimulus figure (i.e., the divided and undivided part aligned on a horizontal line). However, the illusory effect is also modulated by the orientation of the figure in the viewing area due to directional anisotropy of the visual field (Bertulis & Bulatov, 2005; Bulatov & Bertulis, 1999).

EXPLANATIONS AND MODELS

Hering (1861), and independently Kundt (1863), starting with a speculative hypothesis on sensory representation of perceived extents, tried to explain the illusion as an effect of illusory shortening of the unfilled extent due to the retina's curvature ("secants theory"). Aubert (1865) questioned Kundt's theory but did not give an alternative explanation. Hering and Kundt's theory was attacked by Helmholtz (1867) on more ideological grounds, as a paradigmatically "nativist" theory of perception. Helmholtz (1867) himself did not offer any better explanation than that "in all sensory perception, distinctly recognizable differences appear to be larger than less distinct differences of the same objective magnitude" (p. 562)—a general observation rather than a developed theory.

Wundt (1897) proposed eye movements and fatigue of oculomotor muscles as a possible cause of the illusion, a theory later found untenable (Spiegel, 1937). However, the idea of continuous visual scanning of the environment as a basis for the illusory effect seems to be resurrected in Piaget's (1969) "centration theory," only replacing Wundt's physiological cause (oculomotor effort) by psychological abstraction (distribution of attention). The attention distribution theory seems to be supported by Piaget and Bang's (1961) data as well as Vurpillot's (1973) studies on OKI-like figures embedded in real-world scene drawings. Yet it is disputable whether these findings represent the core mechanism of the illusory effect or whether they reflect only secondary, modulating factors.

In the spirit of Gestalt psychology, Spiegel (1937) interpreted the OKI as a result of attractive "*Z*-force" acting between elements of similar appearance in the visual field and counteracted by repulsive forces between interpolated elements. (A third kind of interaction, an "*X*-force" was added to account for differences between dark and white background stimuli.) In this theory, the effect is due to the empty part shrinking relatively to the unfilled part, similarly as in the Hering–Kundt theory but based on a different set of hypotheses. Spiegel suggested that the organizing "forces" act in the central (cortical) representation, rather than on a peripheral (retinal) level. In a similar vein, Eriksson (1970) presented a rather general model of geometric-optical illusions, including the OKI, based on attractive/repulsive interactions between visual elements.

Tausch (1954), in his theory of geometric-optical illusions, sought to explain the OKI by unconscious inferences from alleged distance cues, arguing that higher density of interspersed elements indicates a texture observed from a larger distance. This theory—a predecessor of Gregory's (1963) popular "inappropriate constancy scaling" theory—is unconvincing: arguably, Figures II.35-1a through II.35-1f

do not offer any real-world interpretation, neither in terms of perspectival nor textural cues. Moreover, Tausch's and related speculative theories fail to account for the most stable feature of the illusory effect—that is, its nonmonotonicity (illusion magnitude decreasing with higher filling densities).

Later work aimed more at plausible models of the illusory effect than at general theories of illusory phenomena. Erdfelder and Faul (1994) presented an information integration model that is, in a way, a generalization of the Hering–Kundt theory. Bulatov et al. (1997) elaborated a more neurophysiologically oriented model of the OKI (and of some other illusions of extent, e.g., Müller-Lyer) based on spatial filtering of the image in the visual system and showing very good agreement with experimental data. Wackermann and Kastner (2010) proposed a functional form modeling the dependences of the effect magnitude on both major determinants of the OKI (i.e., on the number of subdividing elements and their vertical extent). Their model does not offer a causal explanation but may be a step toward a phenomenological theory connecting different variants of the illusory expansion phenomenon (Wackermann, 2010).

RELATED PHENOMENA

Dimensional distortions of quadratic shapes filled with horizontal or vertical line patterns (Helmholtz, 1867), or of rectangular shapes sectioned with varied number of lines (Botti, 1909), can be considered a special variant of the OKI. An OKI-like interaction with the Poggendorf illusion was described by Weintraub and Tong (1974). In addition to these unidirectional effects observed in two-dimensional figures, there are analogies to genuinely two-dimensional effects. Obonai (1933, reported in Robinson, 1998, p. 52) suggested a relation between the OKI and Delbœuf's illusion observed in concentric circles. Giora and Gori (2010) studied filled space expansion in squares filled with regular or random patterns and found a nonmonotonic dependence of the filled effects magnitude on numerosity of the filling elements (i.e., the inverse of the dominant spatial frequency). Similar nonmonotonic functional dependences of the effect on visual density of superimposed elements are also known for other types of geometric-optical illusions, such as the Ehrenstein–Orbison figure (Yoshino et al., 2009). A three-dimensional variant of the OKI in perception of spatial arrangements is also known (Deręgowski & McGeorge, 2006).

As for nonvisual modalities, analogies between visual and haptic perception of spatial extents have been reported (Jaensch, 1906; Révész, 1934; Suzuki & Arashida, 1992; cf. Sanders & Kappers, 2009, for a recent review). There is also a parallelism between the filled space illusion and overestimation (i.e., subjective expansion) of filled or fractionated acoustically perceived time intervals (Buffardi, 1971; Grimm, 1934; Thomas & Brown, 1974). Importantly, these findings suggest that the illusory distortion is not specific to sensory systems as such but may be located on higher levels of "information processing." If so, studies of the OKI and similar phenomena would belong rather to the psychophysics of *sensus communis* than to visual psychophysics.

CONCLUSIONS

The OKI has a somewhat unique position among other geometric-optical illusions: misnamed, often misconceived, and still minimally understood. Its fascination is primarily due to its simplicity: as a pure illusion of extent, the OKI clearly demonstrates the dependence of the visual field metric on its content (Wackermann & Kastner, 2009; Westheimer, 2008). Another, no less puzzling aspect of the OKI is the interplay of factors codetermining the illusory effect.

REFERENCES

Aubert, H. (1865). *Physiologie der Netzhaut*. Breslau, Germany: Morgenstern.

Bertulis, A., & Bulatov, A. (2005). Distortions in length perception: Visual field anisotropy and geometrical illusions. *Neuroscience and Behavioral Physiology, 35*, 423–434.

Boring, E. G. (1942). *Sensation and perception in the history of experimental psychology*. New York: Appleton-Century-Crofts.

Botti, L. (1906). Ein Beitrag zur Kenntnis der variabeln geometrisch-optisch Streckentäuschungen. *Archiv gesamter Psychologie*, *6*, 306–315.

Buffardi, L. (1971). Factors affecting the filled-duration illusion in the auditory, tactual, and visual modalities. *Perception & Psychophysics, 10*, 292–294.

Bulatov, A., & Bertulis, A. (1999). Distortions of length perception. *Biological Cybernetics*, *90*, 185–193.

Bulatov, A., Bertulis, A., & Mickienė, L. (1997). Geometrical illusions: Study and modelling. *Biological Cybernetics*, *77,* 395–406.

Coren, S., & Girgus, J. S. (1978). *Seeing is deceiving: The psychology of visual illusions*. Hillsdale, NJ: Lawrence Erlbaum.

Deręgowski, J. B., & McGeorge, P. (2006). Oppel–Kundt illusion in three-dimensional space. *Perception, 35*, 1307–1314.

Doricchi, F., Guariglia, P., Figliozzi, F., Silvetti, M., Gasparini, M., Merola, S., … Bueti, D. (2008). No reversal of the Oppel–Kundt illusion with short stimuli: Confutation of the space anisometry interpretation of neglect and "cross-over" in line bisection. *Brain, 131*, 1–4.

Dworkin, L., & Bross, M. (1998). Brightness contrast and exposure time effects on the Oppel–Kundt illusion. *Perception*, *27*(ECVP Abstract Suppl.).

Erdfelder, E., & Faul, F. (1994). Eine Klasse von Informations-Integrations-Modellen zur Oppel–Kundt-Täuschung. *Zeitschrift für Psychologie, 202*, 133–160.

Eriksson, E. S. (1970). A field theory of visual illusions. *British Journal of Psychology, 61*, 451–466.

Giora, E., & Gori S. (2010). The perceptual expansion of a filled area depends on textural characteristics. *Vision Research, 50*, 2466–2475.

Gregory, R. L. (1963). Distortion of visual space as inappropriate constancy scaling. *Nature, 199*, 678–680.

Grimm, K. (1934). Der Einfluss der Zeitform auf die Wahrnehmung der Zeitdauer. *Zeitschrift für Psychologie, 132*, 104–132.

Jaensch, E. (1906). Über Täuschungen des Tastsinns. *Zeitschrift für Psychologie, 41*, 280–294.

Helmholtz, H. (1867). *Handbuch der physiologischen Optik*. Leipzig, Germany: Barth.

Hering, E. (1861). *Beiträge zur Physiologie*. Vol. 1. Leipzig, Germany: Engelmann.

Knox, H. W., & Watanabe, R. (1894). On the quantitative determination of an optical illusion. *The American Journal of Psychology, 6*, 413–421, 509–514.

Kundt, A. (1863). Untersuchungen über Augenmaass und optische Täuschungen. *Annalen der Physik und Chemie, 196*, 118–158.
Lewis, E. O. (1912). The illusion of filled space. *British Journal of Psychology, 5*, 36–50.
Luckiesh, M. (1922). *Visual illusions, their causes, characteristics and applications.* New York: Van Nostrand.
Messer, H. (1876). Notiz über die Vergleichung von Distanzen nach Augenmaas. *Annalen der Physik und Chemie, 233*, 172–175.
Obonai, T. (1933). Contributions to the study of psychophysical induction: III. Experiments in the illusions of filled space. *Japanese Journal of Psychology*, 699–720.
Oppel, J. J. (1854–1855). Ueber geometrisch-optische Täuschungen. *Jahresberichte des physikalischen Vereins zu Frankfurt,* 37–47.
Oppel, J. J. (1860–1861). Ueber geometrisch-optische Täuschungen (Zweite Nachlese). *Jahresberichte des physikalischen Vereins zu Frankfurt,* 26–37.
Piaget, J. (1969). *The mechanisms of perception.* New York: Basic Books.
Piaget, J., & Bang, V. (1961). L'évolution de l'illusion des espaces divisés (Oppel–Kundt) en présentation tachistoscopique. *Archives de Psychologie, 38*, 1–21.
Piaget, J., & Osterrieth, P. A. (1953). Recherches sur le développement des perceptions XVII : L'évolution de l'illusion d'Oppel–Kundt en fonction de l'âge. *Archives de Psychologie, 34*, 1–38.
Révész, G. (1934). System der optischen und haptischen Raumtäuschungen. *Zeitschrift für Psychologie, 131*, 295–375.
Ricci, R., Pia, L., & Gindri, P. (2004). Effects of illusory spatial anisometry in unilateral neglect. *Experimental Brain Research, 154*, 226–237.
Robinson, J. O. (1998). *The psychology of visual illusion* (2d ed.). Mineola, NY: Dover.
Rothwell, B., & Zaidel, E. (1990). Visual field differences in the magnitude of the Oppel–Kundt illusion vary with processing time. *Perception & Psychophysics, 47*, 180–190.
Sanders, A. F., & Kappers, A. M. L. (2009). Factors affecting the haptic filled-space illusion for dynamic touch. *Experimental Brain Research, 192*, 717–722.
Sanford, E. C. (1903). *A course in experimental psychology.* Boston: Heath.
Savazzi, S., Frigo, C., & Minuto, D. (2004). Anisometry of space representation in neglect dyslexia. *Cognitive Brain Research, 19*, 209–218.
Savazzi, S., Posteraro, L., Veronesi, G., & Mancini, F. (2007). Rightward and leftward bisection biases in spatial neglect: Two sides of the same coin? *Brain, 130*, 2070–2084.
Spiegel, H. G. (1937). Über den Einfluß des Zwischenfeldes auf gesehene Abstände. *Psychologische Forschung, 21*, 327–383.
Surkys, T., & Bertulis, A. (2010). Oppel–Kundt illusion weakens with exposure shortening. In A. Lukoševičius & A. Kriščiukaitis (Eds.), *Proceedings of the International Conference on Biomedical Engineering* (pp. 57–60). Kaunas, Lithuania: Technologija.
Suzuki, K., & Arashida, R. (1992). Geometrical haptic illusions revisited: Haptic illusions compared with visual illusions. *Perception & Psychophysics, 52*, 329–335.
Tausch, R. (1954). Optische Täuschungen als artifizielle Effekte der Gestaltungsprozesse von Größen- und Formenkonstanz in der natürlichen Raumwahrnehmung. *Psychologische Forschung, 24*, 299–348.
Thomas, E. A. C., & Brown, I. (1974). Time perception and the filled-duration illusion. *Perception & Psychophysics, 16*, 449–458.
Turner, R. S. (1994). *In the eye's mind: Vision and the Helmholtz–Hering controversy.* Princeton, NJ: Princeton University Press.
Vurpillot, E. (1963). *L'organisation perceptive: son rôle dans l'évolution des illusion optico-géométriques.* Paris: Vrin.
Wackermann, J. (2010). Geometric–optical illusions: A pedestrian's view of the phenomenal landscape. In A. Bastianelli & G. Vidotto (Eds.), *Fechner Day 2010* (pp. 171–176). Padova, Italy: International Society for Psychophysics.
Wackermann J. (2012a). Determinants of filled/empty optical illusion: Differential effects of patterning. *Acta Neurobiologiae Experimentalis, 72*, 89–94.
Wackermann, J. (2012b). Determinants of filled/empty optical illusion: Influence of luminance contrast and polarity. *Acta Neurobiologiae Experimentalis, 72*, 412–420.
Wackermann, J., & Kastner, K. (2009). Paradoxical form of filled/empty optical illusion. *Acta Neurobiologiae Experimentalis, 69*, 560–563.
Wackermann, J., & Kastner, K. (2010). Determinants of filled/empty optical illusion: Search for the locus of maximal effect. *Acta Neurobiologiae Experimentalis, 70*, 99–110.
Weintraub, D. J., & Tong, L. (1974). Assessing Poggendorf effects via collinearity, perpendicularity, parallelism, and Oppel (distance) experiments. *Perception & Psychophysics, 16*, 213–221.
Westheimer, G. (2008). Illusions in the spatial sense of the eye: Geometrical–optical illusions and the neural representation of space. *Vision Research, 48*, 2128–2142.
Wundt, W. (1897). Die geometrisch-optischen Täuschungen. *Abhandlungen der Königlichen Sächsischen Gesellschaft der Wissenschaften, 42*, 55–178.
Yoshino, D., Kimura, A., & Noguchi, K. (2009). Visual illusion and aesthetic preference: Some common stimulus properties. *Gestalt Theory, 31*, 29–41.

Chapter 36
The Shifted-Chessboard Pattern as Paradigm of the Exegesis of Geometrical-Optical Illusions

Gerald Westheimer

BEGINNING OF THE QUEST

A visual percept that differs from its generating object may make the philosopher ponder the tenuous link between the content of consciousness and reality, but when an effect is so widely prevalent and repeatable to be categorized as an illusion, it will prompt the vision scientist's interest.

Geometrical-optical illusions, in which quite primitive geometrical attributes such as position, length, orientation, or straightness of contours are misperceived in a black-and-white configuration of the simplest kind, emerged as a topic in mid-19th century (Westheimer, 2008). The origin and direction of the intellectual currents that gave birth to it deserve to be put into proper historical perspective; some factors contributing to that development would surely include the successes of physics and chemistry to define a measurable, lawful world of outside objects; the unfolding of physiology as a scientific discipline; the transition from visual observation to photography in astronomy; the acceptance of observer reports as a scientific datum; and many more.

But the veridicality of size judgment had already been under challenge for a long while before these events. The moon illusion has a venerable history, going back to Aristotle. Early in the 19th century there was a heated debate about the apparent increase in size of objects with brightness (Goethe, 1808). Physiologists (Aubert, 1865; Volkmann, 1863), even astronomers, were drawn in because at the time their observations were all visually based. The diffraction theory of light, according to which all stars are imaged as a small bright disk, was involved, proving that the image of even the smallest object would not be a point. The phenomenon was widely discussed under the heading of "irradiation." But a more sophisticated explanation was demanded by the claim that not only small bright but also small dark objects appeared to increase in diameter with increasing background brightness (Volkmann, 1863).

Explanations that posited the site of the illusions in perceptual processes were counterposed to those that looked to defects in the eye. After all, didn't Helmholtz (1868/1903) famously say that if an optician delivered to him an optical instrument with the many faults of the human eye, he would send it back? (Writing about the human eye, Helmholtz actually wrote: "If an optician wanted to sell me an instrument that had such defects I would feel justified to chastise him severely about the carelessness of his work and return the article under protest." He was quick to add, however: "Concerning my own eyes, of course, I would not do that, but to the contrary be glad to keep them as long as possible in spite of all their faults" [p. 286]).

This dialectic is with us still: Is an illusion the consequence of the physics or physiology of processing visual signals, or should one pursue investigations that seek explanation in the imperatives of perception? Of course to the followers of Hering, this is not a dichotomy: for them the physiology of the visual process cannot be disambiguated from the imperatives of perception.

BOTTOM-UP: VISUAL OBJECT TO "EARLY" VISUAL PROCESSING

The shifted-chessboard illusion described by Münsterberg in 1897 is a particularly good case in point, because it is a rich mine for a bottom-up mode of analysis. In chess or checkerboard patterns where the squares in alternate rows are shifted by only a half of a module, the contour lines between the rows appear sloped rather than parallel (Fig. II.36-1). Other names of the illusion came from the fact that at one time café walls (Gregory & Heard, 1979) were tiled with, and kindergarten blocks (Brosterman, 1997) featured, such patterns (Fig. II.36-2). The effect came on to the scene relatively late in the history of geometrical-optical illusions (Robinson, 1998); Oppel, Zoellner, Hering and even Müller-Lyer preceded it. Of the important ones, only Ponzo's came later. Hence when it appeared, it immediately invited attempts at deconstruction, starting with Pierce's (1898) critique the following year, still current today, and continuing to be topical more than a century later.

The story could, however, not be fully pieced together until the individual constituents, known to be involved all along, had been worked up and characterized quantitatively in the second half of the 20th century. In the sequence in which they operate, as the stimulus makes its

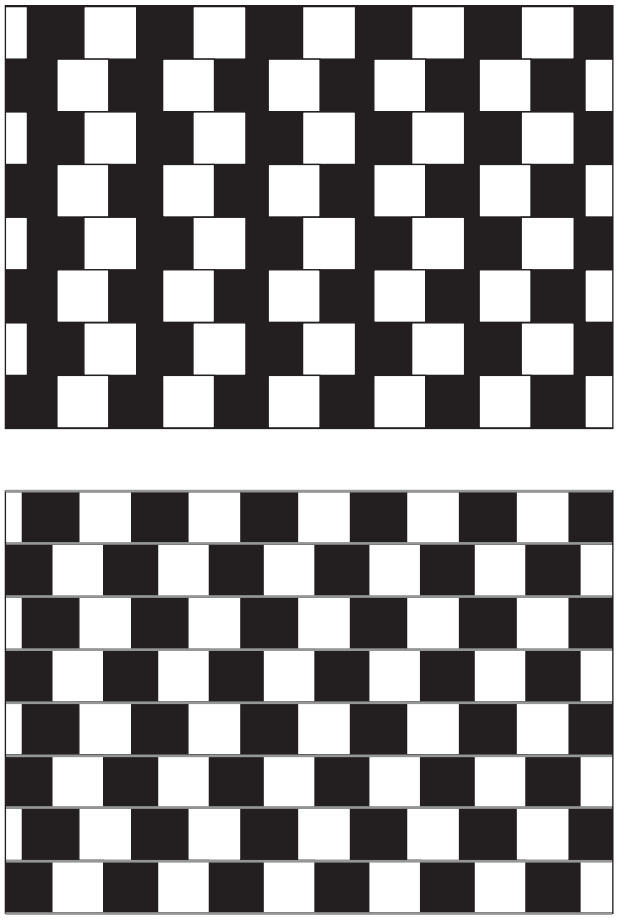


Figure II.36-1. Top: Typical shifted-chessboard pattern. In alternate rows the squares are displaced one-half of square width. Below: Some analyses have introduced "mortar lines" in the shifted chessboard pattern (Morgan & Moulden, 1986). The illusion is robust to many such transformations but not to substitution of isoluminant heterochromatic stimuli for the original black and white (Westheimer, 2007). (Prepared by author.)

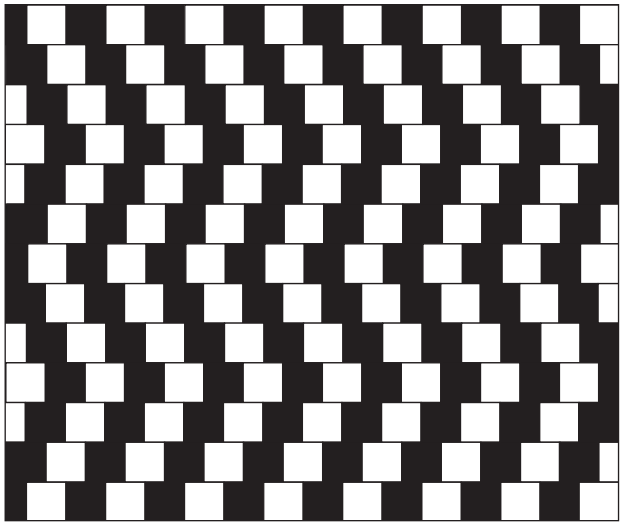


Figure II.36-2. Kindergarten or café wall version of illusion included in the attempt at its deconstruction published by Pierce within a year of Münsterberg's (1897) paper. (Prepared by author.)

way from the object world toward the observer's percept, they are

- the influence of the eye's optics in creating the optical image of the retina.
- neural processing in the retina.
- "early" cortical visual processing.

Itemized in this manner, they indeed form a hierarchical sequence, because neural operations in the retina do not feed back to or interact with the optical image and, in the mammal, the traffic from the retina to the cortex is entirely one-way.

In its essence, the illusion hinges on the apparent orientation assumed by the dividing line between contours when a set of borders is made up of partially offset elements, compared with the rectangularity in a conventional chessboard pattern.

Sufficient information about the transmission properties of the eye's optics is available to provide a serviceable account of the retinal light distribution and compute the convolution of the target light patterns with the eye's point-spread function (Westheimer, 2006). This shows, as was known long ago of course, that there light bleeds from the light to the dark side of an edge (Fig. II.36-3, curve B). But it does not automatically follow, as Helmholtz (1868/1903) pointed out in the 1860s, that subjectively the edge would be localized on the dark side of the border; after all, the point of inflexion of the light distribution remains at the geometrical position of the geometrical edge. More recently we have become aware of some significant properties of the involved retinal neural circuitry. The two most relevant are a compressive nonlinearity in the translation from receptor activity to excitation in the outer region of the retina (Naka & Rushton, 1966) and the spatial center/surround antagonism (Ratliff, 1965). Both are involved in conditioning the

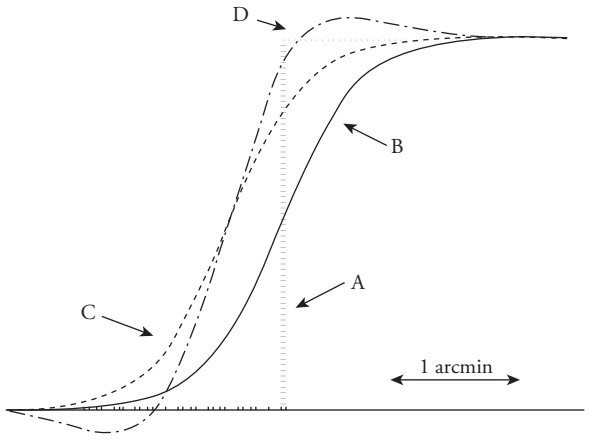


Figure II.36-3. Spatial distribution at an edge. (A) Dotted line: Geometrical position of the dividing line at a black/white edge. (B) Solid line: Spread of light in the retinal image. (C) Dashed line: Excitation level in the outer retina after processing of the light by the receptors and a compressive nonlinearity in the intensity. (D) Dot-dashed line: Pattern of impulses in an array of retinal ganglion cells when an additional center/surround antagonistic reorganization is folded in. Distributions are shown normalized and refer to the situation in the human central fovea. (From Westheimer, 2007.)

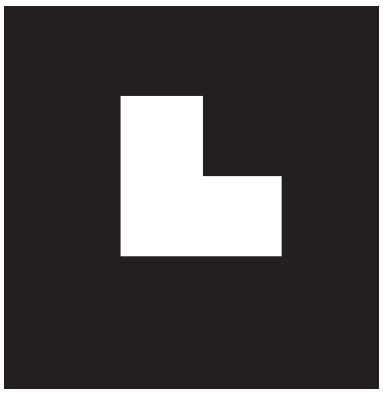


Figure II.36-4. Basic element of the shifted-chessboard figure in object space is the L-shaped component. (Prepared by the author.)

spatial pattern of ganglion cell firings, which constitutes the final output of the retina on which perception is based (Fig. II.36-3d).

What makes the shifted chessboard differ from a regular rectangular pattern and hence responsible for the illusion are the L-shaped constituents (Fig. II.36-4). They are also present in all the variants of the illusion (Pierce, 1901; Wade, 1982). Examining the effect on such a picture element sequentially of the spread of light by the eye's optics, the compressive nonlinearity in the intensity domain, and the spatial center/surround antagonism gives a hint of the rendering of the configuration within the output realm of the ganglion cell fibers (Fig. II.36-5, top). The percept would emerge through assembly of such components that had been subjected to the neural processing in the early cortical stages.

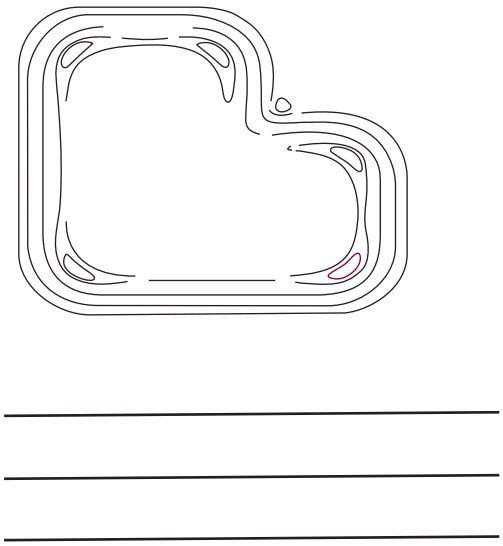


Figure II.36-5. After convolution with the processing mechanism detailed in Figure II.36-3, the L-shaped components assume an outline shown in the top part of the figure. Here the significant element is the slope of the inner edges and the rendering of the interior right angle as acute. This would have the consequence of giving the boundaries between rows a sawtooth configuration (bottom); they are seen, however, as smooth and continuous. (From Westheimer, 2007.)

An effective mode of analysis in such situations was utilized in the pioneering study of Oppel (1854–1855), which he called *Gegenversuch* (literally, "countering experiment") or, as we now call it, nulling. The target is presented with a distorted light distribution designed to cancel those imposed by the hypothesized modification due to the intervening processing apparatus. If the stratagem works, the pattern would now be perceived with strictly rectangular contour lines and the illusion could properly be ascribed to the aberrations in the optical and retinal preprocessing.

This turns out to be only partially successful (Westheimer, 2007). To the extent that it fails, other factors are involved beyond those that had been hypothesized, factors of central neural origin. But as distinct from the separate and sequential consideration of the optical and retinal mechanisms, the central one cannot be dealt with in such a direct manner. Once signals enter the cortex and are shuttled within and between neural circuits, they are subject to interaction from sources both at their own level (e.g., the primary visual cortex) and by feedback from others further along (e.g., top-down influences). Consideration must be given to phenomena such as attention, memory, previous exposure and learning. Here quantitative reasoning based on anatomy or derived from neurophysiological recording in nonprimate and in anesthetized preparations has only limited validity.

TOP-DOWN: PERCEPTUAL IMPERATIVES

Trimming mechanistic elements originating in the peripheral sensory apparatus allows a more unimpeded focus on geometrical illusions as indicators of perceptual channeling (Gillam, 1980). For example, suppose it is granted that the interior right angle of the L-shaped tokens that make up the shifted-chessboard illusion is rendered acute by the preprocessing factors of the eye's optics and of the neural retina. This by itself would give the dividing line a sawtooth contour (Fig. II.36-5, lower), whereas it is universally seen to be smooth, continuous, and possessing a single orientation. We are obviously dealing with a perceptual operation of a different kind, namely, the merging of a run of discontinuous segments into a continuous border with a single orientation. This is not the same problem as the one of assigning an orientation to a short line, which is claimed as one of the great successes of a psychophysical → neurophysiological convergence. Even small offsets in a sequence of line segments lead to a prominent reduction in its orientation attribute (Westheimer & Ley, 1997), hence much remains to be elucidated about the neural apparatus generating the percept of smooth continuous dividing lines with orientations differing by only a very few degrees.

Going a whole step further: How might one approach this particular geometrical-optical illusion as the consequence of its interpretation as an object configuration that gives the appearance of alternating converging and diverging borders? In other words: What three-dimensional disposition of what targets in the observer's object space would yield such a percept? For example, the shifted chessboard pattern in Figure II.36-1 might be thought of as representing a fence made of horizontal slats with rectangular markings, with slats pointing forward at the right alternating

with slats pointing back. This kind of an analysis is not pursued here, but presumably it would need inquiry into the way the markings on adjacent slats would have to be arranged to make dividing lines appear to converge toward the right or the left edges of the figure. Proponents of a strongly phenomenal examination may not deem it relevant to extend such an analysis into the realm of the neural substrate that might yield such a transformation. But, if pursued, it would lead to the question of the manner by which a higher-order internal interpretative apparatus imposes a lack of parallelism on a neural representation that initially passed through the peripheral sensory stages as parallel. Top-down influences on the excitation patterns of units in the primary visual cortex, which would be involved in such operations, are the subject of current investigations.

Recent discussions have looked to Bayesian inference as a possible analytical tool in perception (Knill & Richards, 1996). As applied here, this would take the following form: the retinal image of a pattern with offset markings in alternate rows might, with a certain probability, arise from a physical structure with alternate rows crisscrossing in space, pointing toward the observer on the right and left in alternation. The likelihood of this image/object association would need to be specified numerically, within the ensemble of possible objects. Further, the prior probabilities of the presence of this range of objects also needs to be known. With this information, the posterior probability can be computed that the observer will respond to the presentation of the pattern in accord with the illusion. In spite of enthusiastic sponsorship of the program of perception as Bayesian inference, it awaits convincing realization.

CONCLUSIONS

This chapter contains a reasonably full account of the bottom-up analysis of a geometrical illusion. Readers are invited to compare the advances made since 1898 when, only a year after Münsterberg's (1897) paper, A. H. Pierce (1898) published his deconstruction containing most of the elements presented here. Advances in psychophysical research in the second half of the 20th century have allowed better numerical characterization of the involved factors. They have also provided a better opportunity for the technique defined in Oppel's very first paper of geometrical-optical illusions as *Gegenversuch*, the measuring of the effect by deliberately introducing deformations to cancel the imperfections and artifacts due to early optical and neural stages of vision and make the target appear true. In this way, the influence of central visual processing can be better laid bare. When, as is the case with the shifted-chessboard illusion, this does not provide a full account of the visual illusion, it at least helps to convey some hints of the principles of perceptual processing that the geometrical-optical illusions betray and that make them such perennial favorites.

REFERENCES

Aubert, H. (1865). *Physiologie der Netzhaut*. Breslau, Morgenstern.

Brosterman, N. (1997). *Inventing kindergarten*. New York: Abrams.

Gillam, B. (1980). Geometrical Illusion. *Scientific American*, *242*(1), 86–95.

Goethe, J. W. (1808). *Zur Farbenlehre: Didaktischer teil*.

Gregory, R. L., & Heard, P. (1979). Border locking and the café wall Illusion. *Perception*, *8*, 365–380.

Helmholtz, H. (1903). Die neueren Fortschritte in der Theorie des Sehens. In H. Helmholtz, *Vorträge und Reden* (Vol. 1, pp. 265–365). Braunschweig: Vieweg. (Original work published 1868)

Knill, D.C., & Richards, W. (1996). *Perception as Bayesian inference*. Cambridge, UK: Cambridge University Press.

Morgan, M. J., & Moulden, B. (1986). The Münsterberg figure and twisted cords. *Vision Research*, *26*, 1793–1800.

Münsterberg, H. (1897). Die verschobene Schachbrettfigur. *Zeitschrift für Psychologie*, *15*, 184–188.

Naka, K. I., & Rushton, W. A. H. (1966). S-potentials from colour units in the retina of the fish *(Ciprinidae)*. *Journal of Physiology*, *185*, 536–555.

Oppel, J. J. (1854–1855). Ueber geometrisch-optische Täuschungen. *Jahresbericht des physikalischen Vereins zu Frankfurt am Mains*, 39–47.

Pierce, A. H. (1898). The illusions of the kindergarten patterns. *Psychological Review*, *5*(3), 233–253.

Pierce, A. H. (1901). *Studies in auditory and visual space perception*. London: Longmans Green.

Ratliff, F. (1965). *Mach bands*. San Francisco, CA: Holden-Day.

Robinson, J. O. (1998). *The psychology of visual illusion*. Mineola, NY: Dover.

Volkmann, A. W. (1863). *Physiologische Untersuchungen im Gebiete der Optik*. Leipzig: Breitkopf und Härtel.

Wade, N. (1982). *The art and science of visual illusions*. London: Routledge.

Westheimer, G. (2006). Specifying and controlling the optical image on the retina. *Progress in Retinal and Eye Research*, *25*, 19–42.

Westheimer, G. (2007). Irradiation, border location and the shifted-chessboard pattern. *Perception*, *36*, 483–494.

Westheimer, G. (2008). Geometrical-optical illusions and the neural representation of space. *Vision Research*, *48*, 2128–2142.

Westheimer, G., & Ley, E. J. (1997). Spatial and temporal integration of signals in foveal line orientation. *Journal of Neurophysiology*, *77*(5), 2677–2684.

PART III
Brightness/Lightness/Color

Chapter 37
A Layered Experience of Lightness and Color

Barton L. Anderson

INTRODUCTION

When we look around the world, we perceive a three-dimensional arrangement of surfaces, materials, and objects that are embedded in a particular illumination environment. Our visual system seems to effortlessly sort the different contributions of image variation into their environmental causes. We can easily tell the difference between light surfaces in low illumination from dark surfaces in high illumination. Variations in surface pigmentation are easily distinguished from luminance variations caused by the three-dimensional shape of an object, unless the variations in pigment were crafted by an artist or other image-generating device. Specular reflections are experienced as what they are—reflections of the illumination field—rather than intrinsic variations in the reflectance properties of a glossy object. The light-transmitting properties of translucent and transparent materials and surfaces can also be distinguished from other forms of surface optics. These different surface and material properties are responsible for the structure in light that we used to identify objects and materials, which provides information used to guide our behavior and choices.

One of the oldest and most extensively studied aspects of surface perception is surface lightness. Lightness (or albedo) refers to the proportion of light a surface reflects. There are a number of problems in understanding how the visual system recovers lightness from images. The amount of light reflected from an opaque, matte surface depends on its lightness (reflectance), the strength and position(s) of the illuminant, and the orientation of the surface relative to the light source. This means that the amount of light that reaches the eye is a confounded mixture of its orientation, reflectance, and illumination environment. In order to recover lightness, the visual system has to somehow "discount" the effects of the surface's orientation and illumination environment from the luminance a surface projects to the eye. One way of thinking of this problem is that the visual system "peels off" the effects of illumination to extract surface reflectance, and represents illumination and reflectance as separate causal layers. So-called "layers" models of lightness and color perception assert that the visual system treats the illuminant as a separate causal "layer" that overlies the intrinsic surface property of lightness (Adelson, 1993; Anderson, 1997, 2003b; Anderson & Winawer, 2005, 2008; Bergstrom, 1977; Gilchrist, 1977).

One of the difficulties in assessing whether the visual system decomposes images into layers to recover surface lightness is that it is not phenomenologically clear whether we experience an impression of both surface lightness and illumination intensity when we view a surface. Moreover, from a mathematical perspective, the problem of disentangling the illuminant from surface reflectance problem seems hopelessly intractable and underconstrained. Illumination and reflectance coexist at every point in the image, and there is no obvious source of information that would specify either the prevailing illumination or the true surface reflectance. These issues have led numerous authors to abandon layers models of lightness perception in favor of a variety of alternative models (Adelson, 1999; Gilchrist et al., 1999).

One way to pursue the issue of whether layered image decompositions play a role in the perception of lightness is to use displays that evoke explicit percepts of layers, such as occurs during the perception of transparency. The perception of transparency involves a decomposition of single luminance value (or color) into a two distinct sources: a near, transparent layer and an underlying surface visible through it. The physics of simple forms of transparent layers (i.e., transparent materials that do not scatter or refract light) is similar to the physics of illumination, which suggests that the two might be processed very similarly and/or involve similar or identical mechanisms (Adelson, 1999; Gerbino, 1994; Gerbino, Stultiens, Troost, & de Weert, 1990; Metelli, 1985). Both transparent filters and shadows (multiplicatively) reduce the amount of light returned from a surface. Transparent surfaces also typically reflect additional light from the transparent layer. More complex translucent materials scatter and/or refract light, blurring and distorting surfaces seen through the translucent object's body (Faul & Ekroll, 2011; Fleming, Jakel, & Maloney, 2011; Singh & Anderson, 2002).

My early work into this problem grew out of my research in binocular occlusion (Anderson, 1994, 1998; Anderson & Julesz, 1995; Anderson & Nakayama, 1994). In addition to its relationship to lightness, transparency is also a form of partial occlusion: the amount of light a transparent surface transmits can be thought of as a measure of its "hiding

power." A variety of previous studies had shown that vivid percepts of illusory transparency could be induced by stereoscopically manipulating the relative depth of a variety of simple displays, such as stereo variants of Kanizsa's illusory contour stimuli (Anderson, 1998; Anderson & Julesz, 1995; Nakayama, Shimojo, & Ramachandran, 1990). These stereoscopic manipulations caused local inducing elements to appear to perceptually split—or *scission*—into multiple layers. Whereas previous work had focused on the relationship these phenomena had to neon color spreading, I discovered that perceived lightness (and/or color) could also be significantly affected by this decomposition (Anderson, 1997, 1998). This can be seen in the stereoscopic images presented in Figure III.37-1. When the two left images are cross-fused (or the two right divergently fused), they appear as an illusory diamond in front of four disks on a white (or black) background. The luminance of the grey inducing elements are identical in the top and bottom pairs of images, but the one on the dark surround looks significantly darker than the one on the light surround. A smaller difference can be observed monocularly, or when the diamond appears to complete amodally behind the inducing elements (by cross-fusing the two right figures or divergently fusing the two left figures). Similar effects can also be experienced in color (Fig. III.37-2): the figures on the light surround appear very light, almost glowing, whereas the bottom two figures in each pair appear as dark filters.

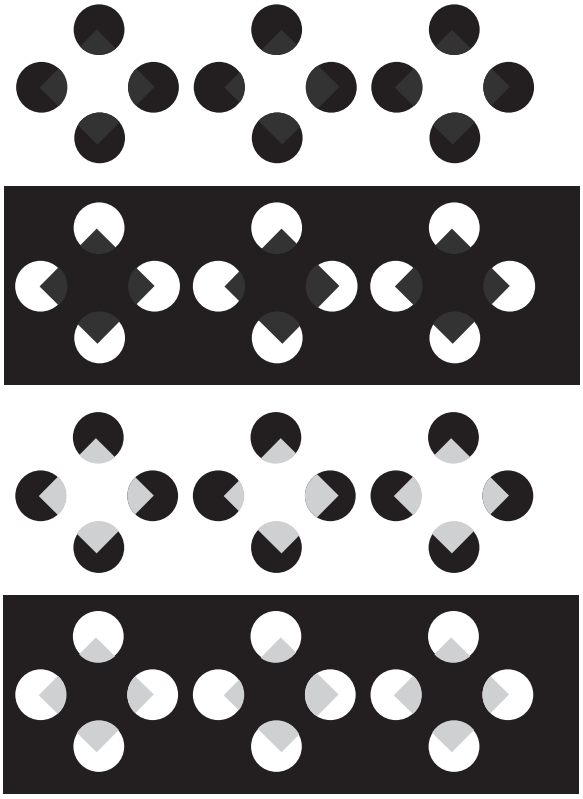


Figure III.37-1. The effects of stereoscopic depth and surround colors on perceived lightness. All of the grey pie-shaped segments in the top and bottom two rows are identical shades of grey. When the figure completes modally in front of the surround and inducers, the apparent lightness difference is quite large, whereas it is smaller when the inducers complete amodally to form a diamond behind the surround and inducers. The Kanizsa figures on the light surrounds (first and third rows) appear quite light, whereas those on the dark surrounds (second and fourth rows) appear quite dark. (Adapted from Anderson, 1998.)

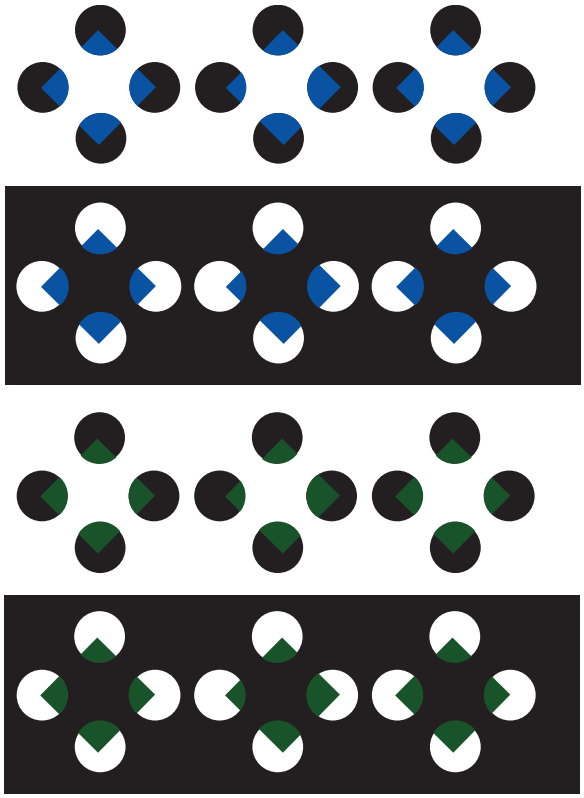


Figure III.37-2. Same as Figure III.37-1 but with colored inducing segments. (Adapted from Anderson, 1998.)

The similarity of these experiences to those experienced in neon color spreading inspired me to suggest that all of these phenomena (and many phenomena that are similarly constructed and lead to similar percepts, such as White's effect) arise from the same cause: the scission of a target into multiple layers (Anderson, 1997, 2001, 2003c). The direction of the transformation in the perceived lightness or color can be predicted by the composition of the multiple layers. For the figures in the top pairs of Figures III.37-1 and III.37-2 (light background), the target sectors are decomposed into a light filter over dark (black) disks, whereas the figures in the bottom of each pair appear as dark transparent surfaces overlying light, white disks. In these displays, the near layer is the "figure" whose lightness and/or color is judged; note that the opposite pattern of induction would be reported if observers were required to report the color of the underlying disks (e.g., the regions underlying the transparent layers for the white Kanizsa inducers look much lighter, essentially "white," while the transparent diamond figure appears much darker). As with most areas of science, and perhaps particularly in lightness perception, this view was not widely embraced. Part of the problem was phenomenological: some people assert that they do not experience a greater illusion in the near versus far or nonstereoscopic case. I therefore set out to construct a version where the experience of layers, and the effect that scission had on lightness, was more explicit.

Some examples of these effects can be seen in Figures III.37-3 and III.37-4 (Anderson, 1999, 2003b). The stimuli in Figure III.37-3 depict a sinusoidal luminance modulation (a grating) placed within a diamond-shaped aperture. The phase of the sinusoid is horizontally shifted in the apertures shown to each eye, which generates a binocular disparity between the aperture boundaries and the grating. When the two images on the right are cross-fused (or the two on the right divergently fused), the grating appears as a coherent surface behind the edges of the aperture. However, when the disparity relationships are inverted (cross-fusing the two on the left or divergently fusing the two on the right), an entirely different percept emerges. The grating now appears to split into two components, which depend critically on their luminance relationships with the background: the grating on the black background (a) appears as a white diamond that is partially occluded by fuzzy black stripes, whereas the grating in (b) appears as a black diamond partially occluded by fuzzy white stripes. Importantly, the gratings on the *grey* background appear very unstable and perceptually incoherent. The reason for this is that they fail to satisfy the photometric constraints required for the perception of transparency (see later discussion). Finally, we showed that these effects are not limited to one-dimensional luminance modulations. Similar effects can be observed in the stereoscopic images in Figure III.37-4. The method of construction is similar to Figure III.37-3, except that noise "clouds" are used instead of sinusoidal gratings, which are viewed through three small circular apertures. When fused in the same manner as the preceding figures, the "near" clouds appear to split into two components: dark clouds in front of light disks (top figure) and light clouds in front of dark disks (bottom).

When the preceding work was first published, the discussion of these phenomena focused on their impact on extant theories of stereopsis. The theoretical emphasis stressed that the dramatic transformations in perceived depth that occurred in these displays—particularly with the textured stereograms in Figures III.37-3 and III.37-4—demonstrated the need to reconceptualize the relationship between image disparities and perceived depth (Anderson, 1999, 2003b). These phenomena showed that perceived depth, lightness, and surface opacity are all intrinsically coupled. The disparity relationships in Figures III.37-3 and III.37-4 are identical in all of these images, yet the patterns of depth, lightness, and transparency are dramatically different. These phenomena illustrated that the scission that occurs during the perception of transparency—a

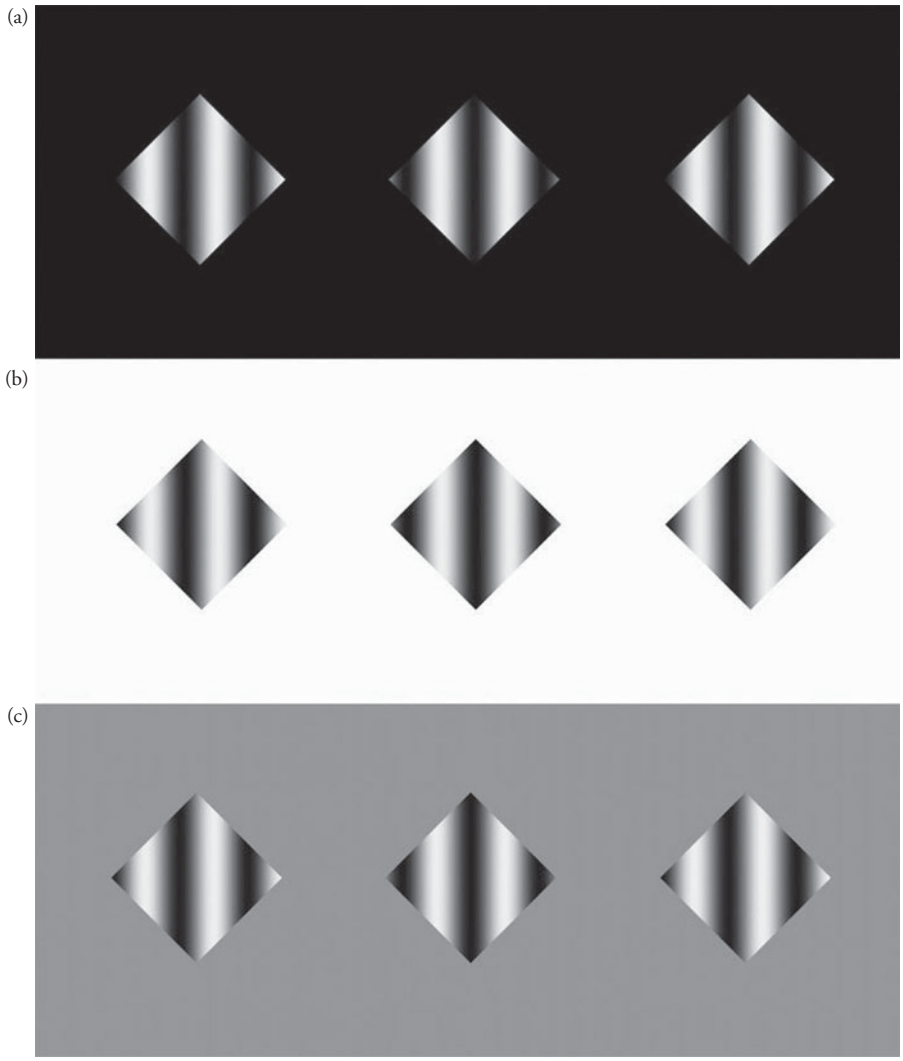


Figure III.37-3. The effects of stereoscopic depth and surround color on the perceived lightness and opacity of one-dimensional textures (sinusoidal gratings). The gratings in all three rows are identical. When the grating is given a disparity that places it in front of the aperture boundary, it appears to split into two layers in (a) and (b), appearing as fuzzy black stripes in front of a white diamond (a) and fuzzy white stripes in front of a black diamond (b). In (c), no coherent percept of surface color is experienced. All of the gratings appear identical as a coherent surface when the disparity is inverted, independently of the surround color. (Adapted from Anderson, 2003b.)

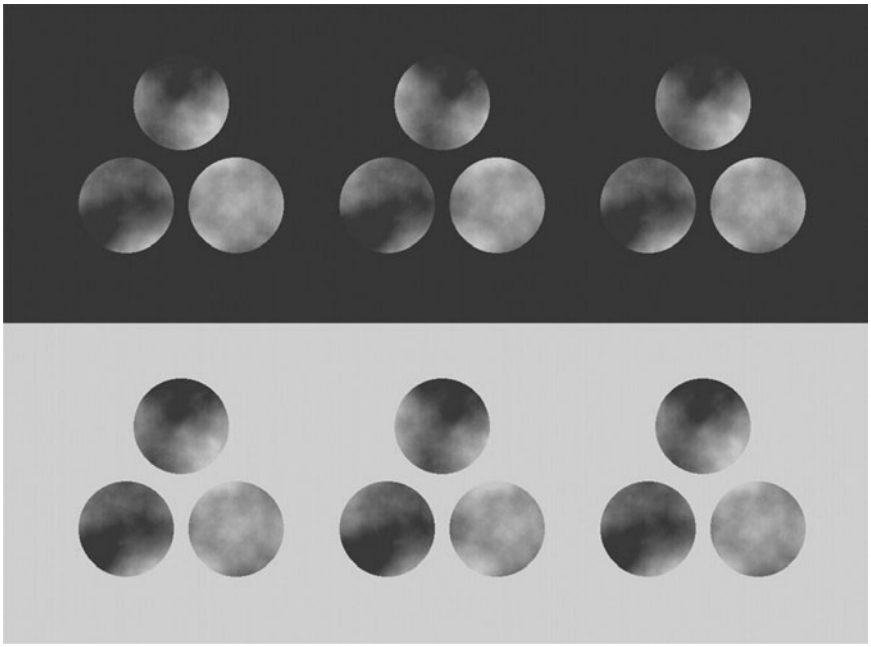


Figure III.37-4. Similar to Figure III.37-3, but the textures now vary in two spatial dimensions. When disparity places the texture in front of the surround, it appears as dark smoke in front of light discs (top) or light smoke in front of dark discs (bottom). (Adapted from Anderson, 2003b.)

decomposition into multiple surfaces—can have a dramatic impact on the perception of lightness, at least in the context of explicit depth differences between different portions of the image.

The next question was whether similar effects could be constructed in the absence of explicit depth differences. The challenge was to create versions of these effects that did not rely on explicit manipulations of the depth of different image components. The idea was to find a method for inducing a scission of a textured target, like that in Figure III.37-4, by causing either the light or dark components of the texture to selectively group with its surround. One of the well-known constraints on perceived transparency is that transparent surfaces cannot reverse the contrast polarity of contours that they overlay. This constraint can be decisive and is responsible for the instability experienced when fusing the bottom of Figure III.37-3. Note that the border of the diamond aperture in this display reverses polarity as the grating goes from light to dark, whereas the two other figures have a consistently signed contrast polarity along their boundaries (dark to light in A, light to dark in B, referenced from surround to grating figure). Only the consistently signed edge supports scission into layers, and the contrast polarity of the edge determines the perceived lightness of the underlying diamond (white for the dark to light contour in A, black for the light to dark contour in B). Thus if we were going to succeed in constructing a monocular version of the illusion, we would have to respect this polarity constraint. In addition, if we were going to attempt to cause only the light or dark "components" within the texture in the target to selectively group with the surround, the texture modulations in the surround would have to be reasonably well "aligned" with the texture variations in the targets. At first, these constraints appeared to be very difficult to simultaneously satisfy for a randomly chosen texture. How could we construct a texture defined by random noise that would just happen to align with the darker and lighter components in its (random) surround? It seemed impossible, at least until we realized that this constraint was already trivially satisfied in the texture image that we used to construct the target.

We began by synthesizing a noise image with an amplitude spectrum of $1/f^2$, and the different frequencies and orientations were randomized and summed, giving us an initial "seed" image texture (Anderson & Winawer, 2005, 2008). We then created three transformed versions of this display: one that varied from black to mid-grey, another that varied from mid-grey to light, and another that simply had its contrast enhanced (see Fig. III.37-5). Note that the ordinal luminance relationships are preserved in all of these transformations (the darkest pixels remain the darkest, and the lightest remain the lightest). This last high-contrast texture was used to construct our targets and was contrast enhanced to produce the maximal lightness transformation possible (we varied the target's contrast in the experiments that we subsequently performed). All we had to do to preserve the textural continuity of the target and

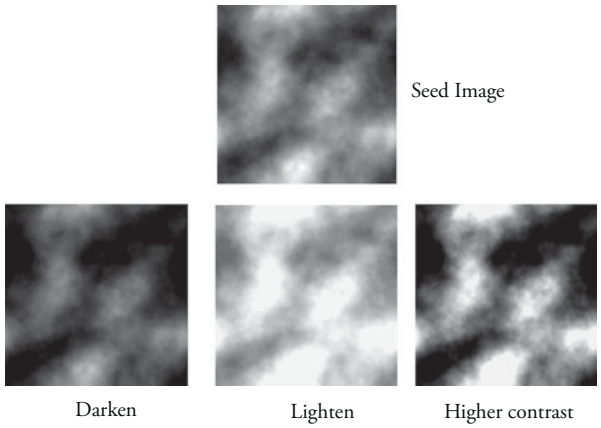


Figure III.37-5. Method used for constructing the images in Figure III.37-6.

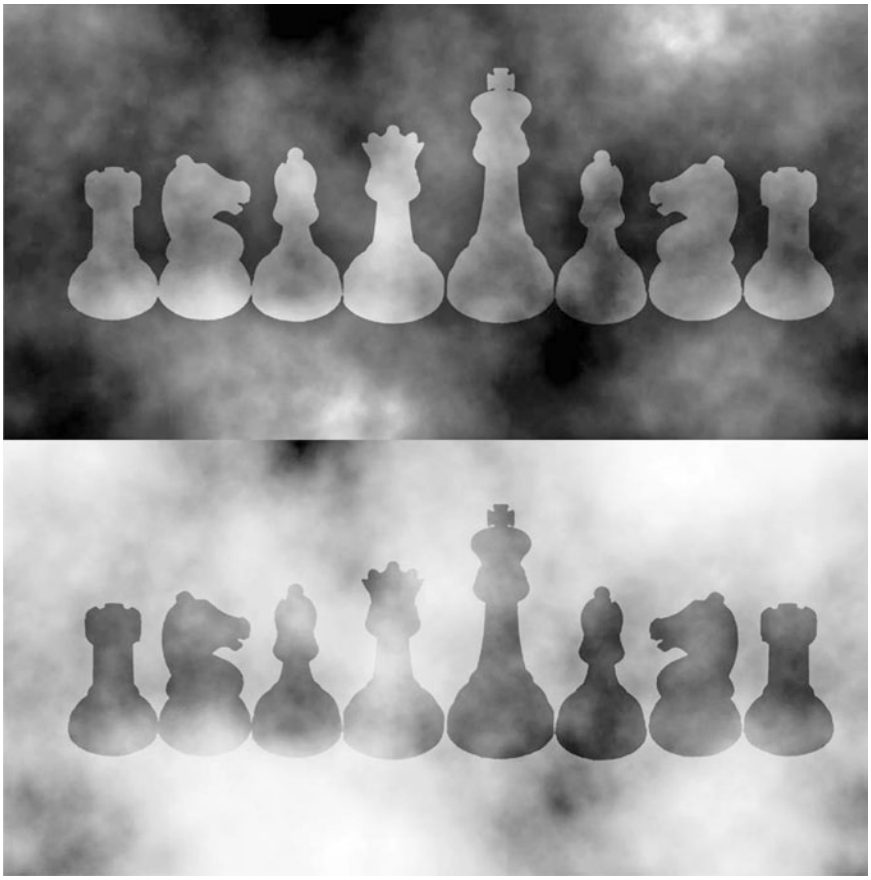


Figure III.37-6. The role of scission on perceived lightness in nonstereoscopic images. The textures inside the chess pieces are identical in the top and bottom rows but appear as white chess pieces behind black smoke (top) and black chess pieces behind with smoke (bottom). (Adapted from Anderson & Winawer, 2005.)

surround was to "cut" our targets from a specified location on the target image and place it on the corresponding position in either of the two surrounds. To insure that the polarity of the target–surround boundary was constant along its length, we simply had to insure that we had sufficiently lightened or darkened the two surrounds. An example of this procedure is illustrated in Figure III.37-6. The effect on perceived lightness is dramatic: the chess pieces in the top of the figure appear white, while the corresponding and physically identical chess pieces in the bottom appear black. Note that the dark components of the texture inside the chess pieces on the top appear to form portions of dark, semitransparent clouds, whereas the opposite is true for the bottom display (i.e., the light components in the targets on the bottom appear as clouds).

I also showed that similar effects could be observed (Fig. III.37-7) in the widely studied and highly debated lightness/brightness illusion known as White's effect (Anderson, 2003a). Although a variety of explanations have been offered for the classic version of this illusion using uniformly colored targets, only scission models are capable of explaining the textured variants of this effect shown in Figure III.37-7. It is noteworthy that both the classical illusion (bottom of Fig. III.37-7) and the textured variants (top of Fig. III.37-7) exhibit a similar dependence in illusion strength on the scale (size) of the stripes used in the display.

As with some of the early demonstrations using variants of neon color spreading displays, similar effects can be observed in the chromatic domain (see Fig. III.37-8; Anderson & Khang, 2010). As with the chess pieces in Figure III.37-6, the disks in the two displays of Figure III.37-8 are physically identical but appear to differ dramatically in color. The consistent contrast polarity of the target–surround border is the critical factor in the lightness displays. In the chromatic variants of these effects, the critical factor needed to induce coherent scission is a consistent direction of chromatic contrast in color space.

THEORETICAL IMPLICATIONS, DEBATES, AND OPEN QUESTIONS

The capacity to generate both chromatic and achromatic transparency displays that exhibit striking transformations in perceived color and lightness by manipulating a target's context suggests that there are common processes that underlie these effects. However, there is some debate about the nature of these processes and their relevance to theories of lightness and color perception more broadly.

One of the motivations for developing textured variants of these displays was to provide clear phenomenological proof that the way that the visual system distributes luminance or chromaticity between the perceptual layers that are experienced in conditions of transparency can have a direct effect on perceived lightness and color. The preceding demonstrations appear to provide incontrovertible evidence in support of this view. However, the relevance of these phenomena for the *general* problem of lightness and color perception, or for understanding other forms of

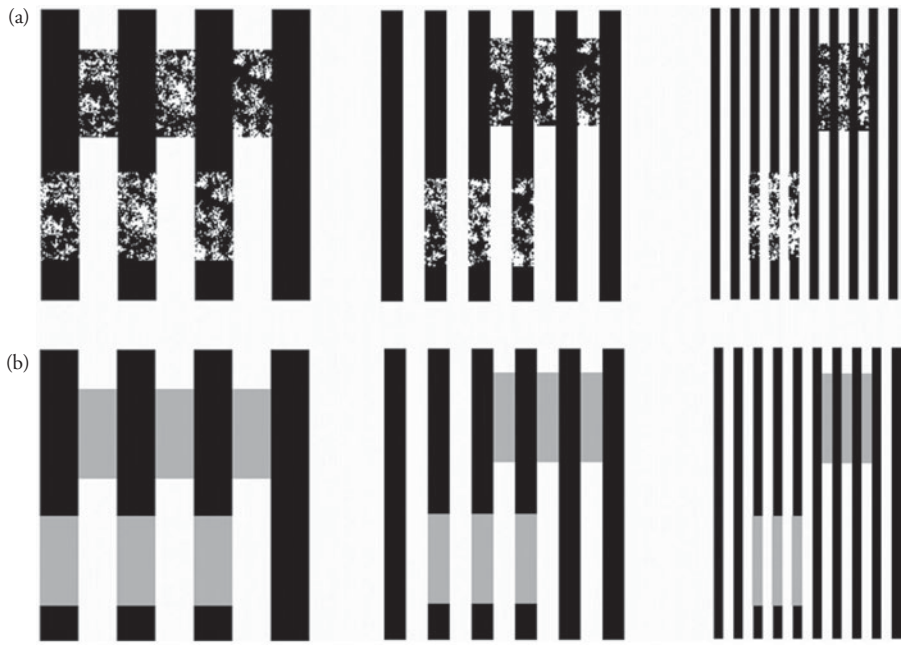


Figure III.37-7. Textured versions of White's effect (top row) are compared with White's effect (bottom row). Note that both the images containing textured targets exhibit a similar dependence on the spatial frequency of the stripes, suggesting that the two effects reflect similar processes. (Adapted from Anderson, 2003a.)

lightness and color effects, remains unclear and is currently debated. There are two general concerns (a) whether the type of scission exhibited in the textured transparency displays is truly a form of scission and (b) whether the problem of "discounting" the illuminant can be construed as a form of scission or layered image decomposition. I consider each issue in turn.

The first issue involves the relationship between transparency and occlusion. As noted previously, transparency is a form of partial occlusion, where a transparent surface's opacity is a measure of its hiding power. Part of the issue revolves around a methodological prejudice: most studies of surface lightness and color use homogenous targets that can be construed as having a particular (and uniform) lightness or albedo. The lightness and color transformations exhibited in Figures III.37-1 and III.37-2 provide examples of the kind of targets that might be used in a traditional experiment on lightness or color. However, the percept of transparency that are demonstrated in Figures III.37-3, III.37-4, III.37-6, and III.37-8 generate percepts of inhomogeneous transparency, where the transparent surface varies in perceived opacity throughout the display. Moreover, in all of these images, there is always a region of the more distant surface (the target layer in these displays) that appears in plain view (i.e., unobscured by the transparent layer). This has prompted the suggestion that these effects are qualitatively

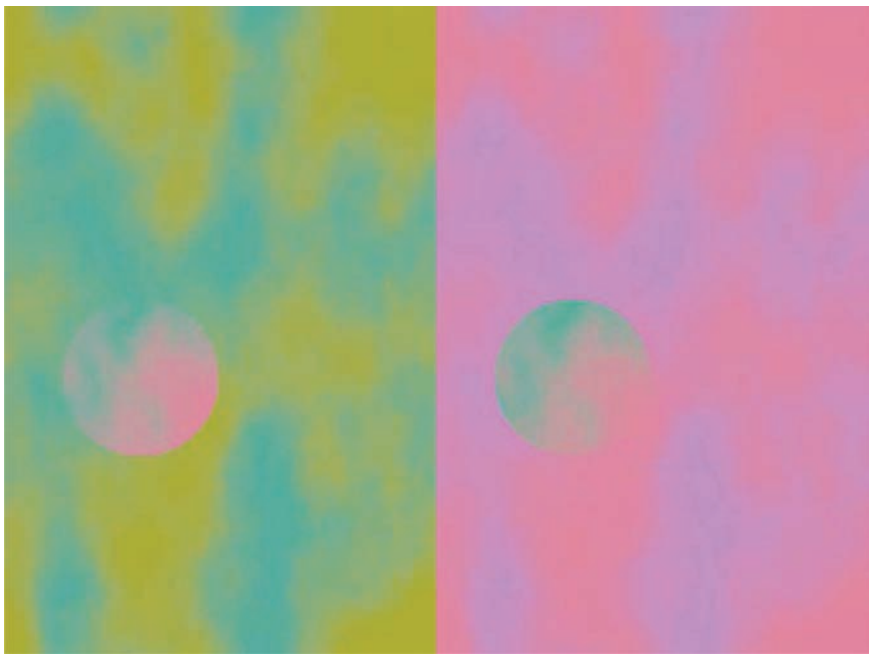


Figure III.37-8. Similar forms of texture scission can be generated by purely chromatic textures. (Adapted from Anderson & Khang, 2010.)

different from other forms of lightness and chromatic induction, which has led some to redescribe these effects as "figure–ground" phenomena (Albert, 2007; Kingdom, 2011). Indeed, there are deep connections to these phenomena and the problem of border ownership (or figure–ground assignment), which were repeatedly discussed during the development of these effects (Anderson, 1996, 1999, 2003b; Anderson, Singh, & Fleming, 2002). Nonetheless, there are a host of problems with arguments that attempt to *reduce* these phenomena to a mere issue of figure–ground assignment. First, for images containing luminance or chromatic gradients, there are no simple edges within the targets to which border ownership can be assigned. The gradients in these displays are experienced as regions of overlay in which the transparent layer varies in opacity. Any sense of a "boundary" between regions in plain view and regions in conditions of transparent overlay is a perceptually *outcome* that must be explained. Indeed, one outstanding issue that requires explanation is why all of the luminance and/or chromatic variations in these displays are experienced as variations in the transparent overlay rather than as variations in the perceived reflectance of the underlying layer. Any sense of figure–ground is a perceptual outcome, not an explanation of these effects. Moreover, attempts to reduce these phenomena to figure–ground reversals also overlook the fact that the same transformations in color and lightness can be observed in chromatic and achromatic noise textures than are completely devoid of the local edges needed to support figure–ground computations (Fig. III.37-9; Anderson & Winawer, 2008; Wollschläger & Anderson, 2009).

The second issue (regarding the relevance of these transparency phenomena to the problem of lightness and color perception more generally) is an ongoing research problem and unresolved issue. As noted earlier, the physics of transparent overlays (or at least some simplified version of these physics) and the physics of illumination are strongly related. The amount of light returned by a surface is a multiplicative product of the illuminant and surface reflectance. A transparent surface has two effects on a surface: it multiplicatively attenuates the amount of light returned to the eye (like sunglasses), and it produces an additional source of light from the reflection of light off of the pigment within the transparent surface. Two general issues arise in attempting to understand the relationship between the decomposition that accompanies the perception of transparency and the perception of lightness and color. One concerns whether the processes that generate percepts of transparency are causally responsible for transformations in lightness and/or color that accompany such percepts. The illusions presented herein suggest that the scission that underlies percepts of transparency *is* causally responsible for at least some forms of lightness and color induction, but this does not imply that *all* forms of chromatic and achromatic induction that are accompanied by percepts of transparency are the consequence of scission (see, e.g., Anderson & Khang, 2011). Several authors (including the editors of this volume) have argued that there are numerous lightness illusions that evoke vivid percepts of transparency for which scission plays no explicit causal role in generating the illusions. These issues are far from resolved, but they highlight the fact that the experience of scission is not sufficient to demonstrate its causal role in the perception of surface color and lightness.

The problem is further complicated with displays that do not evoke percepts of transparency. Whereas the perception of layers is explicit in conditions of transparency, it is unclear whether the illuminant is explicitly represented as a distinct causal "layer." Although there is good evidence that we can judge regions that differ in their degree of illumination in natural scenes, it is unclear how this information is derived and represented. The fact that we can derive reasonably good estimates of surface lightness and color suggests that there is information that allows us to either implicitly or explicitly represent (or "discount") the illuminant, but we are far from a point of consensual

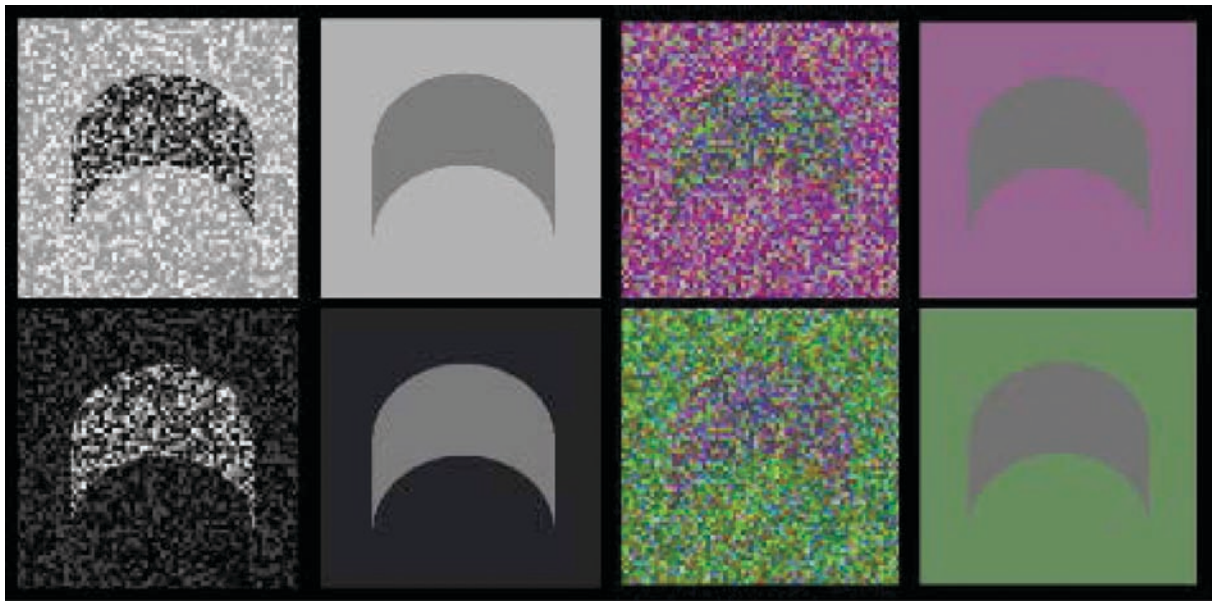


Figure III.37-9. Chromatic induction induced by chromatic scission. Note that there are no edges that can be assigned a figure–ground reversal in these images, particularly the random-dot version of these displays. (Adapted from Wollschläger & Anderson, 2009.)

understanding of how this is accomplished. As with most exciting areas of research, we remain in a condition of having more questions than resolved answers.

REFERENCES

Adelson, E. H. (1993). Perceptual organization and the judgment of brightness. *Science*, *262*(5142), 2042–2044.

Adelson, E. H. (1999). Lightness perception and lightness illusions. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (2d ed., pp. 339–351). Cambridge, MA: MIT Press.

Albert, M. K. (2007). Occlusion, transparency, and lightness. *Vision Research*, *47*(24), 3061–3069. doi: 10.1016/j.visres.2007.06.004

Anderson, B. L. (1994). The role of partial occlusion in stereopsis. *Nature*, *367*(6461), 365–368.

Anderson, B. L. (1996). Border ownership as a special case of contrast ownership. *Perception*, *25*(Abstract Suppl.).

Anderson, B. L. (1997). A theory of illusory lightness and transparency in monocular and binocular images: The role of contour junctions. *Perception*, *26*(4), 419–453.

Anderson, B. L. (1998). Stereovision: Beyond disparity computations. *Trends in Cognitive Sciences*, *2*, 222–228.

Anderson, B. L. (1999). Surface perception. In R. A. Wilson & F. C. Keil (Eds.), *Encyclopedia of the cognitive sciences* (pp. 816–817). Cambridge, MA: MIT Press.

Anderson, B. L. (2001). Contrasting theories of White's illusion. *Perception*, *30*(12), 1499–1501.

Anderson, B. L. (2003a). Perceptual organization and White's illusion. *Perception*, *32*(3), 269–284.

Anderson, B. L. (2003b). The role of occlusion in the perception of depth, lightness, and opacity. *Psychological Review*, *110*(4), 785–801.

Anderson, B. L. (2003c). The role of perceptual organization in White's illusion. *Perception*, *32*, 269–284.

Anderson, B. L., & Julesz, B. (1995). A theoretical analysis of illusory contour formation in stereopsis. *Psychological Review*, *102*, 705–743.

Anderson, B. L., & Khang, B. G. (2010). The role of scission in the perception of color and opacity. *Journal of Vision*, *10*(5), 1–16.

Anderson, B. L., & Nakayama, K. (1994). Toward a general theory of stereopsis: Binocular matching, occluding contours, and fusion. *Psychological Review*, *101*, 414–445.

Anderson, B. L., Singh, M., & Fleming, R. W. (2002). The interpolation of object and surface structure. *Cognitive Psychology*, *44*(2), 148–190.

Anderson, B. L., & Winawer, J. (2005). Image segmentation and lightness perception. *Nature*, *434*(7029), 79–83.

Anderson, B. L., & Winawer, J. (2008). Layered image representations and the computation of surface lightness. *Journal of Vision*, *8*(7), 18.11–18.22.

Bergstrom, S. S. (1977). Common and relative components of reflected light as information about illumination, color, and 3-dimensional form of objects. *Scandinavian Journal of Psychology*, *18*(3), 180–186.

Faul, F., & Ekroll, V. (2011). On the filter approach to perceptual transparency. *Journal of Vision*, *11*(7), 1–33.

Fleming, R. W., Jakel, F., & Maloney, L. T. (2011). Visual perception of thick transparent materials. *Psychologial Science*, *22*(6), 812–820.

Gerbino, W. (1994). Achromatic transparency. In A. L. Gilchrist (Ed.), *Lightness, brightness, and transparency* (pp. 215–255). Hillsdale, NJ: Lawrence Erlbaum.

Gerbino, W., Stultiens, C. I., Troost, J. M., & de Weert, C. M. (1990). Transparent layer constancy. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(1), 3–20.

Gilchrist, A. L. (1977). Perceived lightness depends on perceived spatial arrangement. *Science*, *195*(4274), 185–187.

Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X. J., . . . Economou, E. (1999). An anchoring theory of lightness perception. *Psychological Review*, *106*(4), 795–834.

Kingdom, F. A. (2011). Lightness, brightness and transparency: A quarter century of new ideas, captivating demonstrations and unrelenting controversy. *Vision Research*, *51*(7), 652–673.

Metelli, F. (1985). Stimulation and perception of transparency. *Psychological Research*, *47*(4), 185–202.

Nakayama, K., Shimojo, S., & Ramachandran, V. S. (1990). Transparency: Relation to depth, subjective contours, luminance, and neon color spreading. *Perception*, *19*(4), 497–513.

Singh, M., & Anderson, B. L. (2002). Perceptual assignment of opacity to translucent surfaces: The role of image blur. *Perception*, *31*(5), 531–552.

Wollschläger, D., & Anderson, B. L. (2009). The role of layered scene representations in color appearance. *Current Biology*, *19*, 430–435.

Chapter 38
Color and Luminance

Afterimages, Combinations, and Flicker

Stuart Anstis

AFTERIMAGES

Color and luminance interact in many ways in the human visual system. For instance, the colors in an afterimage, which are due to adaptation of retinal cones, are especially vivid when test contours, presented after the adapting image, coincide with the blurred edges of the afterimage. Daw (1962) first discovered that colored afterimages were much more visible if they were superimposed on a congruent luminance-defined test pattern. A fine example is demonstrated in Sadowski's well-known "Spanish castle illusion" (see www.johnsadowski.com/big_spanish_castle.php).

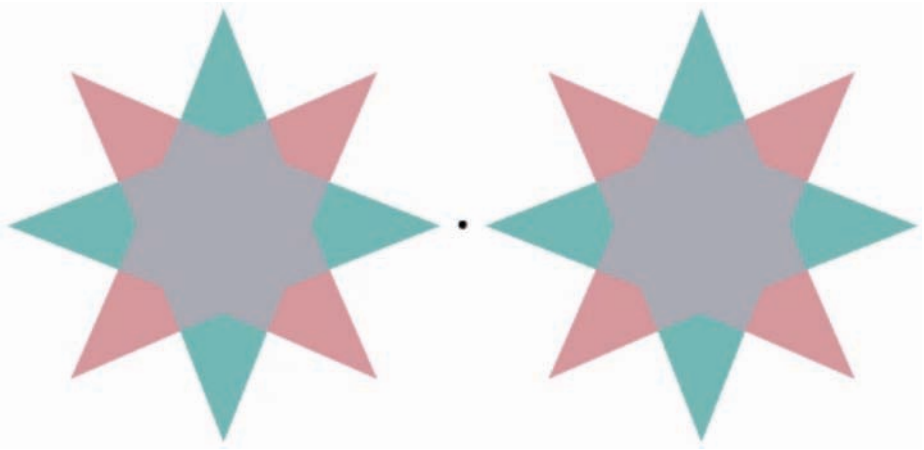
Van Lier, Vergeer, and Anstis (2009) showed that the same adapting pattern can induce multiple, differently colored afterimages at test locations that were not adapted to color (Video III.38-1). This was because afterimage colors spread out and average spatially within test contours but not across them.

The adapting field in Video III.38-2 is a plaid composed of a horizontal red/green grating transparently superimposed on a vertical blue/yellow grating. Following adaptation to this plaid, if one inspects a plain white test field, one sees a conventional negative afterimage. But if the test field contains black horizontal lines, congruent with the edges of the adapting grating, the afterimage consists of a red/green grating. If the test field contains vertical lines, the afterimage is a blue/yellow grating.

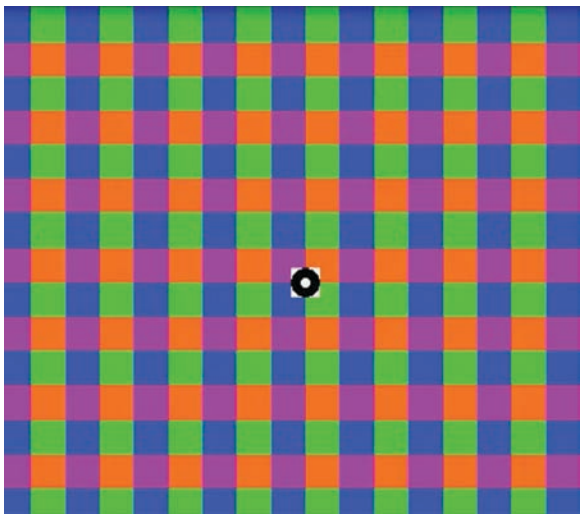
Video III.38-3 shows that second-order test contours made of moving stripes can also constrain the spread of color just as effectively as the luminance test contours in Video III.38-2. Thus the observed filling-in of afterimage colors strongly depends on the test contours presented after the colored stimulus, revealing color–contour interactions that resemble filling-in of "real" colors. This suggests that afterimage colors tend to spread perceptually in a diffusion-like manner, until they encounter a luminance contour that arrests the spread (Grossberg, 2003; Kim & Francis, 2011).

LUMINANCE GATES COLOR: LOOKING AT TWO PAINTINGS AT ONCE

Luminance signals are produced by the sum of the R, G, B retinal cones and color signals by the differences between these cones. Luminance and color pathways have very different spatial properties. Acuity is much higher for luminance contours than for color contours, and it seems that colors diffuse out until they meet a luminance contour (Grossberg, 2003; Kim & Francis, 2011). Thus if

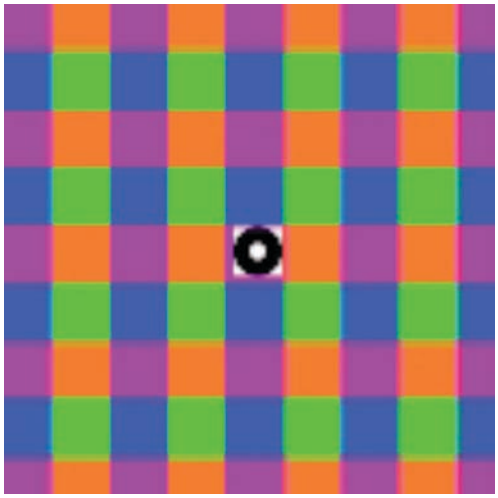


Video III.38-1. A partly colored eight-pointed star is followed by two differently oriented four-pointed stars. The four-pointed star outline that includes the previously colored green points appears to be filled in with a red afterimage color, whereas the subsequently presented four-pointed star outline that includes the red points appears to be filled in with a green afterimage color. That is, the perceived colors of the afterimage depend on the geometry of black/white test contours. (Image by Rob Van Lier.)



Video III.38-2. A single adapting pattern can give very different colored afterimages, depending on the test field. The adapting plaid was made by superimposing a vertical blue/yellow grating on a horizontal red/green grating. A uniform white test field would simply show a negative afterimage. But the thin black vertical lines in the test field enhance the blue/yellow vertical components of the afterimage, while the black horizontal test lines enhance the red/green horizontal components of the afterimage. (Image by Stuart Anstis.)

a colored picture is split into its luminance and color components, and the color component is blurred before the components are recombined, then the picture still looks much the same (Wandell, 1995). This difference can be brought out by a new *two-picture* effect (Anstis, Vergeer & Van Lier 2012a, b), which demonstrates how a grayscale picture can gate or select a fuzzy chromatic picture in the presence of noise. In short, two paintings, O1 and O2, are split into their luminance (greyscale) components L1 and L2 and their color components C1 and C2. The two color components C1 and C2 are transparently superimposed.



Video III.38-3. Second-order test bars made of moving achromatic random dots also induce the colored afterimages. (Image by Stuart Anstis.)

Adding the greyscale of the first painting (= C1 + C2 + L1) looks like the original O1, while adding the greyscale of the second painting (= C1 + C2 + L2) looks like the original O2. In conclusion, the luminance contours select or gate the congruent–color contours and ignore non-congruent colors from the other painting. Two full-length portraits of about the same size and shape are superimposed in this example: *The Blue Boy* (1770), by Thomas Gainsborough (1727–1788), Huntington Art Gallery, San Marino, California; and *La Source* (1856), by Jean-Auguste-Dominique Ingres (1780–1867), Musee d'Orsay, Paris. These paintings have very different colorings. *The Blue Boy* is in cool bluish tones while *La Source* is infused with warm flesh tones. Each painting is first split into its luminance (greyscale) component and its chrominance (color) component, and each component is saved as a separate file (a web tutorial on this technique is given by John Sadowksi at http://www.johnsadowski.com/color_illusion_tutorial.html).

Of course simply superimposing the luminance and chrominance components of a single painting would restore the original. But instead, just the color components of *both* paintings are superimposed in Figure III.38-1. On its own, this looks like a fuzzy mess. But Figure III.38-1 shows two

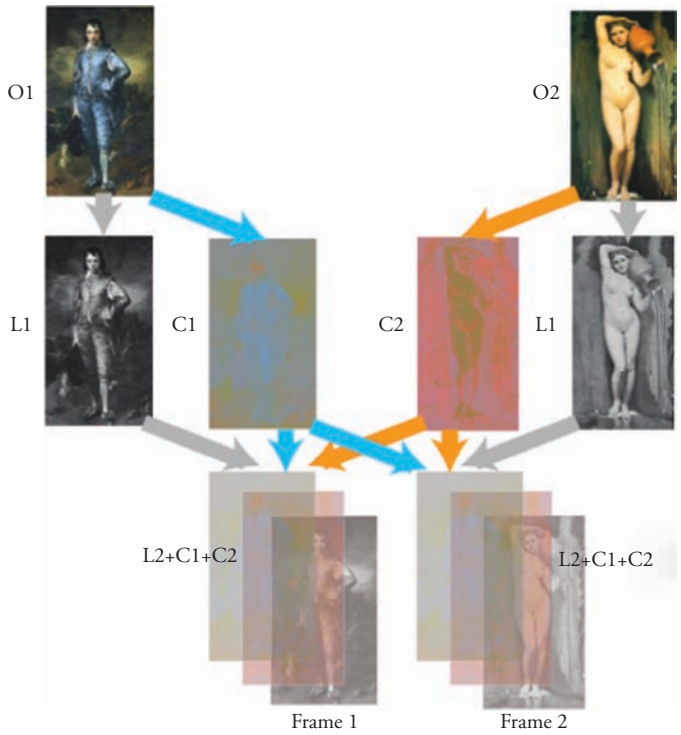


Figure III.38-1. The two original paintings, *The Blue Boy* (O1) and *La Source* (O2) are split into their luminance components (L1, L2) and their color components (C1, C2) (second row). These are recombined, and transparently superimposed, in a two-frame movie (third row). Frame 1 contains *both* color pictures plus L1, and frame 2 contains *both* color pictures plus L2. Result: The luminance pictures dominate and gate the colors, so Frame 1 looks like *The Blue Boy*, while Frame 2 looks like *La Source* (see also Video III.38-1). (Thomas Gainsborough, *The Blue Boy*, Huntington Art Gallery, San Marino, California.)



Figure III.38-2. (a) Contains C1 + C2 + L1 and looks like O1 (*The Blue Boy*). (b) Contains C1 + C2 + L2 and looks like O2 (*La Source*). Conventions same as for Figure III.38-1. (Jean-Auguste-Dominique Ingres, *La Source*, Musee d'Orsay, Paris.)

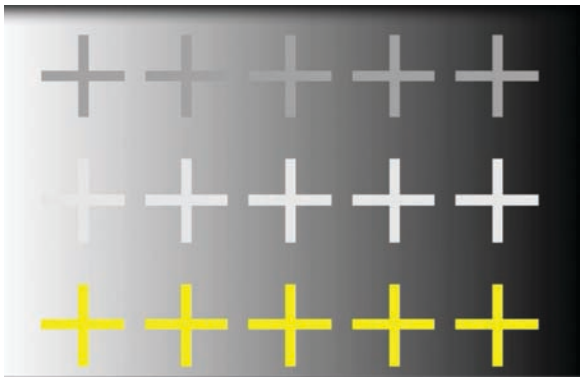
copies of this double-chromatic mess. On one is transparently superimposed the luminance component of *The Blue Boy* and on the other the luminance component of *La Source.*

Thus each frame is a triple-layer sandwich of two chromatic pictures (same in both) and one luminance picture (different in each). These two frames are shown enlarged in Figure III.38-2. Although each frame is somewhat desaturated by a veil of color, *The Blue Boy* luminance picture makes the whole image look like the original *Blue Boy* painting, with *La Source* virtually invisible. Conversely, the original *La Source* painting is perceived when the triple-layer sandwich contains the *La Source* luminance picture, with *The Blue Boy* nowhere to be seen. Thus each greyscale picture perceptually amplifies the colors that are congruent with it and de-emphasizes the noncongruent colors. This is like a visual analog of the auditory cocktail party problem. It is as though the luminance contours pick out the colors that coincide with them but ignore noncoincident colors. Color filling-in phenomena like these are thought to occur by means of a contour-based filling-in mechanism (Grossberg, 2003). Thus luminance perception dominates and guides color perception.

FLICKER-AUGMENTED CONTRAST

It is well known that a grey cross looks darker when it is on a light background and lighter when it is on a dark background (Heinemann, 1955). This is illustrated in the top row of Video III.38-4.

Anstis & Ho (1998) noticed an effect that we called "flicker-augmented contrast" (FAC). If the cross flickers between black and white at, say, 8Hz, instead of being a steady grey, the impression of simultaneous contrast



Video III.38-4. Flicker augmented contrast. Top row: Simultaneous contrast. The crosses are all the same grey, but the left-hand ones on the light surround look somewhat darker than the right-hand ones on the dark surround. Middle row: The crosses all flicker together between black and white. Now the simultaneous contrast is much enhanced; the leftmost cross looks almost black and the rightmost cross looks almost white. Bottom row: The crosses all flicker together between dark blue and light yellow. Now the leftmost cross looks dark blue and the rightmost cross looks light yellow. These apparent colors cannot be directly induced by the achromatic surround. Instead, the visual system selects the most salient edges with the highest luminance differences—blue/white on the left and yellow/black on the right. (Image by Stuart Anstis.)

is greatly enhanced, and a flickering cross looks black on a light background and white on a dark background (second row of Video III.38-4). It turns out that FAC does not work by enhancing the lateral inhibition that is thought to underlie simultaneous contrast (Creutzfeldt, 1993), since FAC is not an induction process but a selection process. The visual system selects the more salient edges, that is, those with the higher Michelson contrast, at the expense of the less salient. In the bottom row of Video III.38-4 the crosses are colored, alternating between dark blue and light yellow. Here a cross looks blue on a light background and yellow on a dark background. Clearly the achromatic light and dark backgrounds contain no colors themselves so they cannot induce colors into the cross. Instead, it is the difference in luminance between the cross and the background that determines the salience and drives the FAC selection process. The dark (blue) is more salient against the light background, and the light (yellow) is more salient against the dark background. When a stimulus appears to vary in salience over time, as the crosses do here, it makes sense for the visual system to pick the most salient and contrasty version of the stimulus.

REFERENCES

Anstis, S., & Ho, A. (1998). Nonlinear combination of luminance excursions during flicker, simultaneous contrast, afterimages and binocular fusion. *Vision Research*, *4*, 523–539.

Anstis, S., Vergeer, M., & Van Lier, R. (2012a). Looking at two paintings at once: Luminance edges can gate colors. *i-Perception*, *3*(8), 515–518.

Anstis, S., Vergeer, M., & Van Lier, R. (2012b). Luminance contours can gate afterimage colors and real colors. *Journal of Vision*, *12*(10), 1–13.

Creutzfeldt, O. D. (1993). The neurophysiological correlates of colour induction, colour and brightness contrast. *Progress in Brain Research*, *95*, 45–53.

Grossberg, S. (2003). Filling-in the forms: Surface and boundary interactions in visual cortex. In L. Pessoa & P. de Weerd (Eds.), *Filling-in: From perceptual completion to cortical reorganization* (pp. 13–37). Oxford: Oxford University Press.

Heinemann, E. G. (1955). Simultaneous brightness induction as a function of inducing and test-field luminances. *Journal of Experimental Psychology*, *50*(2), 89–96.

Kim, J., & Francis, G. (2011). Color selection, color capture, and afterimage filling-in. *Journal of Vision*, *11*, 1–18.

Van Lier, R., Vergeer, M., & Anstis, S. (2009). Filling-in afterimage colors between the lines. *Current Biology*, *19*(8), R323–R324.

Chapter 39
The White Effect

Barbara Blakeslee and Mark E. McCourt

In the White effect (White, 1979) gray test patches of identical luminance placed on the black and white bars of a square-wave grating appear different in brightness/lightness (Fig. III.39-1). This illusion has generated intense research interest because, unlike other induction effects such as classical simultaneous contrast, the direction of the brightness/lightness shift is not consistent with an explanation in terms of either contrast (where the brightness/lightness of a test region shifts *away* from that of adjacent regions) or assimilation (where the brightness/lightness of a test region shifts *toward* that of adjacent regions). This means that the brightness/lightness of the test patch does not correlate with the amount of black or white border in contact with it or in its general vicinity. For example, when the test patch is a vertically oriented rectangle sitting on the white bar of a vertical grating, it has two short sides that are in contact with the collinear white bar on which it sits and two longer sides that are in contact with the flanking black bars (see left-hand test patch in Fig. III.39-1a). This configuration describes a test patch that has more extensive border contact with the black flanking bars, yet the gray patch appears darker than an identical gray patch situated on a black bar and flanked by white bars. The effect cannot simply be attributed to assimilation, however, since the direction of the effect is unchanged even if the height of the test patch is reduced until it has more extensive border contact with the collinear white bar (see left-hand test patch in Fig. III.39-1b). In addition, although the magnitude of the White effect increases with increasing spatial frequency (Anstis, 2006; Blakeslee & McCourt, 2004; White, 1979), a finding that is consistent with assimilation, unlike other assimilation effects (DeValois & DeValois, 1988; Helson, 1963; Jameson, 1985; Jameson & Hurvich, 1989; Smith, Jin, & Pokorny, 2001), the illusion does not disappear or reverse to a contrast effect at low spatial frequencies (Blakeslee & McCourt, 2004; White, 1979). From observations of this type, White concluded early on that explanations couched in terms of either contrast or assimilation that depend solely on the relative amounts of black and white surrounding the gray elements cannot explain the effect and that oriented mechanisms must play a causal role. In other words, the effect cannot be explained by a mechanism such as the "low-level" filtering performed by center-surround receptive fields in the retina and lateral geniculate nucleus of cats and primates (Jameson & Hurvich, 1989; Kingdom, 2011; Kingdom & Moulden, 1988).

In response to these findings, several qualitative filtering explanations incorporating oriented filters were introduced to explain the White effect. White (1981) proposed a mechanism he called "pattern-specific inhibition," where elongated cortical filters that had similar preferred orientation and spatial frequency selectivity, and that received their input from adjacent retinal locations, inhibited each other to produce the effect. Similarly, Foley and McCourt (1985) suggested that hypercomplex-like oriented cortical filters with small centers and elongated inhibitory surrounds might be responsible for the effect. Moulden and Kingdom (1989) proposed a "dual mechanism" model in which a local mechanism mediated by circularly symmetric center-surround receptive fields produced a strong signal at the corner intersections of the test patch with the collinear bar. This corner signal was argued to be responsible for the greater weight given the coaxial bar relative to the flanking bar in producing a brightness/lightness contrast effect in the test patch. A second more spatially extensive mechanism, similar to the filters with small centers and elongated surrounds first proposed by Foley and McCourt, was proposed to account for the increase in the size of the brightness/lightness effect with increasing collinear bar height.

Other investigators, however, have taken a very different view and have rejected spatial filtering explanations for brightness/lightness induction altogether in part due to the White effect. These researchers have instead offered "higher-level" explanations of brightness/lightness effects based on junction analysis (Adelson, 2000; Ripamonti & Gerbino, 2001; Todorovic, 1997; Zaidi, Spehar, & Shy, 1997); Gestalt grouping (Bressan & Kramer, 2008; Gilchrist, 1999; Gilchrist et al., 2006); and/or perceptual inferences about depth, shape, illumination, reflectance, and transparency (Anderson, 1997, 2001, 2003; Adelson, 2000; Howe, 2005; Spehar, Clifford, & Agostini, 2002; Spehar, Gilchrist, & Arend, 1995; Taya, Ehrenstein, & Cavonius, 1995). According to many of these approaches it is perceptual organization, such as relative depth relations in the White effect, that influences brightness/lightness perception, where the rule is that surfaces predominantly interact

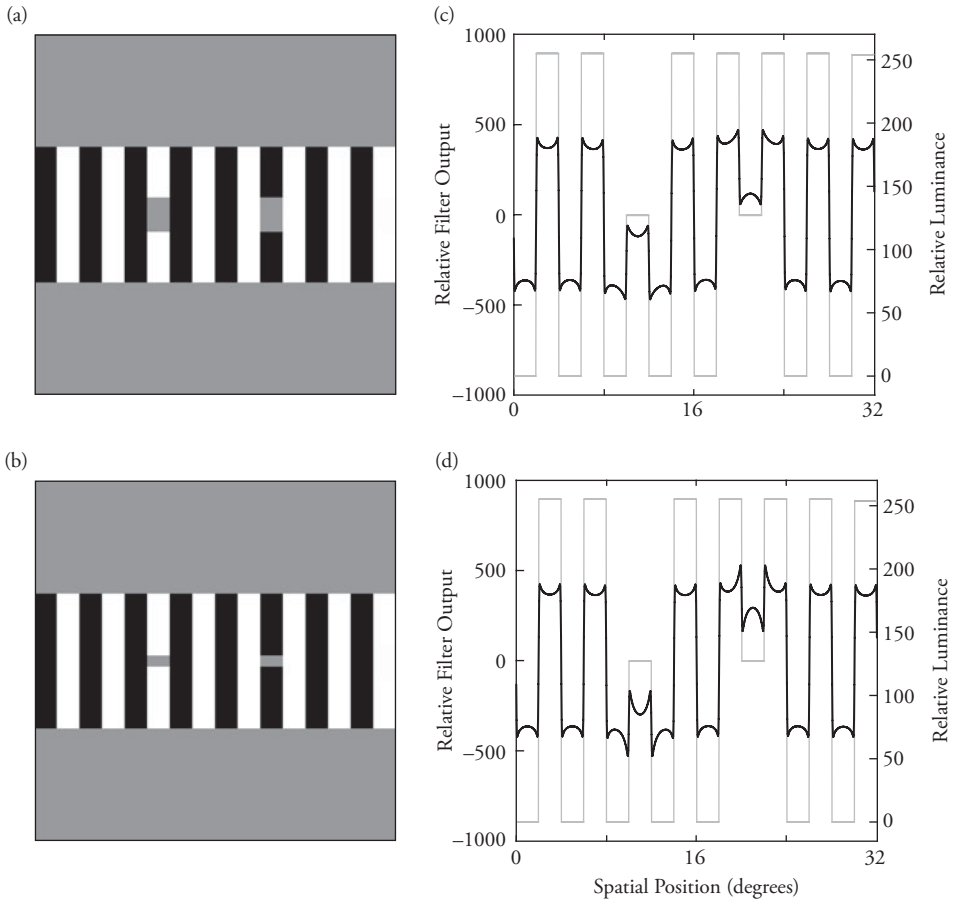


Figure III.39-1. In the White effect gray test patches of identical luminance placed on the black and white bars of a square-wave grating differ in brightness/lightness. In panel (a) the left-hand test patch is a vertically oriented rectangle sitting on the bright/white stripe of a vertical grating; it has two short sides that are in contact with the collinear white bar on which it sits and two longer sides that are in contact with the flanking black bars. This configuration describes a test patch having more extensive contact with the black flanking bars, yet the gray patch appears darker than an identical gray patch situated on a black bar and flanked by white bars (panel [a] right-hand test patch). The White effect cannot simply be attributed to assimilation, however, since the direction of the effect is unchanged even if the height of the test patch is reduced until it has more extensive border contact with the bar on which it is situated (see panel [b]). Figure belongs to B. Blakeslee and M. McCourt, based on data from Blakeslee and McCourt (2004).

with other surfaces with which they are grouped. For example, in the anchoring model of Gilchrist (2006; see also Gilchrist et al., 1999), the lightness of a target in any given perceptual framework is "anchored" by (i.e., assigned relative to) the surface in that framework that appears white, which is usually the surface with the highest luminance. The perceptual frameworks to which a target belongs are defined in terms of Gestalt grouping principles. Thus, in the White effect, the gray test patch appears lighter (or darker) when it is on the black (or white) bar because it "belongs to" or has been "grouped with" that bar. According to Gilchrist et al. (see also Gilchrist, 2006), the principal grouping factor at work here is the T-junction, which they propose signals depth through occlusion. Alternatively, Anderson (1997, 2001, 2003) advocates a "scission" account of the White effect in which the T-junctions in these displays trigger the parsing of the test patches into multiple layers (although not necessarily stratified in depth). Thus, according to this account of the White effect, the test patch collinear with the black bar appears lighter because some of the darkness has been attributed to an underlying black bar. Similarly, the test patch collinear with the white bar appears darker because some of the lightness (white) has been attributed to an underlying white bar.

Interestingly, both Zaidi et al. (1997) and Todorovic (1997) argued that an analysis of local junctions alone, specifically T-junctions, can account for the White effect and that the depth-inducing aspect of T-junctions is not responsible for the effect. Todorovic acknowledged, however, that an earlier study by White and White (1985) revealed a major problem for local T-junction analysis. White and White manipulated the spatial phase (position) of the gray test patches relative to the inducing bars of the White (1979) display. A designation of 0° phase was assigned when the gray test patches were collinear with the black bars of the inducing grating and 180° phase when they were collinear with the white bars. In addition, four other equally spaced phase offsets (36°, 72°, 108°, and 144°) were tested. Discontinuous changes in T-junction structure did not result in discontinuous changes in brightness/lightness; the results instead revealed a nearly linear relationship between the brightness/lightness and the spatial phase of the gray test patches where test patches darkened monotonically as test-patch phase increased from 0° to 180°. Howe (2005) also challenged T-junction explanations of White's effect with a novel variation of the stimulus where the T-junctions remained unchanged but the brightness/lightness effect was either eliminated (for 13 observers) or reversed (for 17 observers). Similarly, Yazdanbakhsh, Arabzadeh, Babadi, and Fazl (2002) showed that the White effect occurs even when the T-junctions at the corners of the test patches are eliminated altogether through the use of ellipsoidal test patches.

Blakeslee and McCourt (1999, 2003, 2004, 2005) and Blakeslee, Pasieka, and McCourt (2005) argued that the rejection of spatial filtering explanations for the White effect (and brightness/lightness phenomena in general) was premature. They provided a mechanistic explanation for the

White effect (and numerous other brightness/lightness effects) in the form of an oriented multiscale filtering model, called the oriented difference of Gaussians (ODOG) model. According to the ODOG model, brightness/lightness induction effects are fundamentally the result of linear spatial filtering with an incomplete basis set. Although the array of spatial filters that comprise the multiscale orientation channels of the ODOG model extend to spatial frequencies much lower than those typically used to model spatial vision in prior models, they nonetheless still form an incomplete set. A complete linear transform (e.g., Fourier or wavelet) is lossless, meaning that a spatial image can be veridically reconstructed from its transform. One obvious consequence imposed by the incompleteness of the filter array of the ODOG model (and by the human visual system it represents) is that arbitrarily high spatial frequency information cannot be represented in a reconstructed image, resulting in spatial smearing (blur). Less well appreciated is the fact that information loss also occurs for low spatial frequencies. Indeed, patterns whose scales are large (i.e., low frequency) relative to the scale of the encoding filters are represented with a loss of low frequency information and therefore exhibit brightness/lightness contrast effects; conversely, patterns whose scales are small (i.e., high frequency) relative to the scale of the encoding filters are represented with a loss of high-frequency information and therefore exhibit brightness/lightness assimilation (Anstis, 2006; Jameson, 1985; Jameson & Hurvich, 1989; Kingdom & Moulden, 1988).

Loss of information due to incompleteness over scale alone was shown to account for effects such as simultaneous contrast, assimilation, and grating induction using the earlier isotropic DOG model of Blakeslee and McCourt (1997). In order to explain these illusions as well as the White effect, however, Blakeslee and McCourt (1999) added orientation selectivity and contrast normalization to the DOG model thus creating the ODOG model (Blakeslee & McCourt, 1999, 2003, 2004; Blakeslee et al., 2005). The

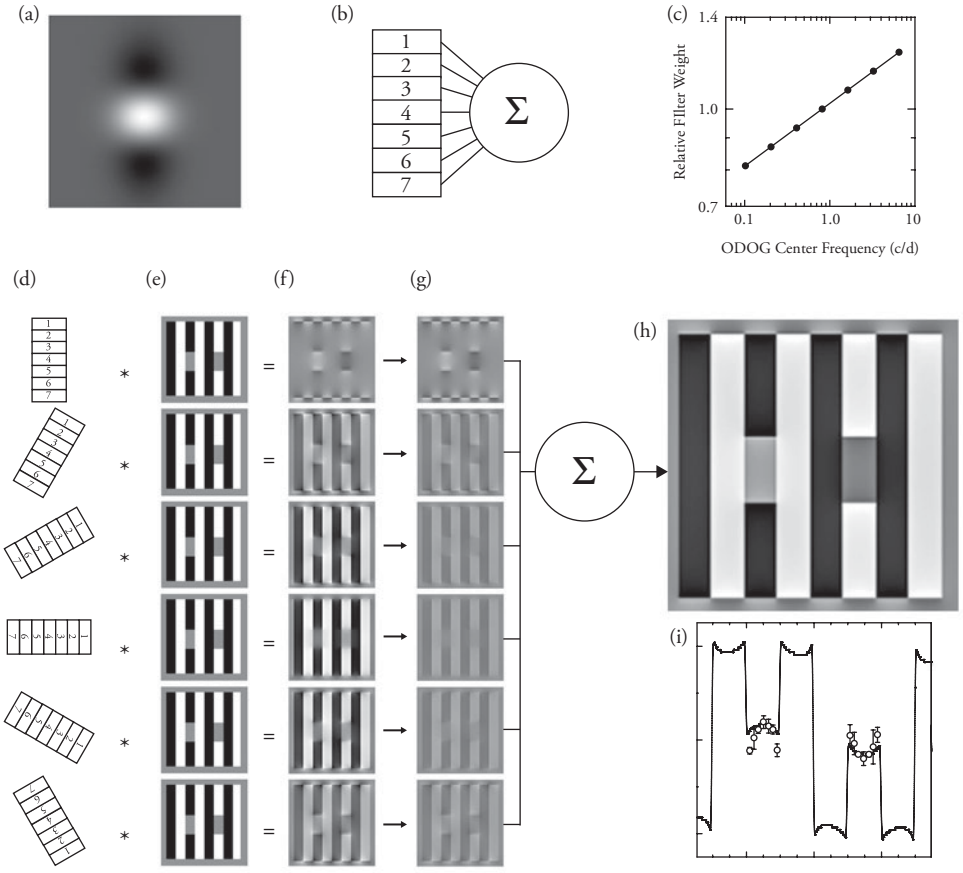


Figure III.39-2. The ODOG model applied to the White effect. Panel (a) illustrates the spatial structure of an ODOG filter. Panel (b) indicates that seven filters at octave-interval spatial scales are summed at each filter orientation. Panel (c) shows the center frequencies of the seven ODOG filters and their relative weighting. Panel (d) illustrates the six channel orientations of the ODOG filters summed across scale. Panel (e) shows the White stimulus. Panel (f) shows the result of convolving each stimulus with the multiscale orientation channels shown in (d). Note the relatively higher level of activity for the horizontally oriented channel. Panel (g) shows RMS normalized activity in each orientation channel. Panel (h) shows the summed output across all six normalized orientation channels. Panel (i) plots a cross-section of ODOG model output to the White stimulus (black line), and point-by-point brightness/lightness matches (open symbols) across the two equiluminant test patches, which demonstrate spatial inhomogeneity. Figure belongs to B. Blakeslee and M. McCourt, based on Figures 1 and 8 in Blakeslee and McCourt (1999).

defining features of the ODOG model (linear multiscale spatial filtering by oriented filters followed by contrast normalization) are characteristics routinely observed at early stages of cortical visual processing in both cats and monkeys (Carandini, Heeger, & Movshon, 1997; Geisler & Albrecht, 1995; Gilbert, Das, Ito, Kapadia, & Westheimer, 1996; Rossi & Paradiso, 1999; Rossi, Rittenhouse, & Paradiso, 1996) and are therefore well justified as components of the model.

The ODOG model is illustrated and summarized in Figure III.39-2. When the long axis of the multiscale ODOG filter is vertical, as it is in the orientation (represented by the top row of Figure III.39-2d–g), the convolution output of the filter with the White stimulus shows the greatest activity in the region of the test patches and accords with the White effect. Although the top and bottom edges of the inducing grating are also a good stimulus for this filter, the inducing grating itself is not. This situation is largely reversed in the convolution output of the multiscale ODOG filter with a horizontal orientation (represented in the fourth row of Fig. III.39-2d–g). Here the activity generated by the inducing grating is high compared to that for the test patches. Contrast normalization prior to summation (Fig. III.39-2g) weights the stimulus features extracted by these two filters equally. This ensures that high-contrast features (e.g., the inducing grating), captured by filters at one orientation, do not overwhelm lower contrast features (e.g., the test patches) captured by filters at other orientations in the combined output (Fig. III.39-2h).

The ODOG model predicts the relative brightness/lightness of the test patches in the White stimulus over a range of spatial frequencies and test patch heights, as well as the magnitude of the White effect relative to other brightness/lightness effects as measured psychophysically using brightness/lightness matching (Blakeslee & McCourt, 1999, 2003, 2004). In addition, the model accounts for the graded change in test-patch brightness/lightness as its spatial position is varied relative to the inducing grating (White & White, 1985), thus disrupting T-junctions, as well as for the variants of the White stimulus described by Howe (2005), in which the T-junctions are preserved but the effect disappears or reverses (Blakeslee & McCourt, 1999; Blakeslee et al., 2005). Finally, the ODOG model also predicts the gradients of brightness/lightness observed across the test patches of White stimuli as measured using point-by-point brightness/lightness matching (Blakeslee & McCourt, 1999; Figures III.39-1c–d and III.39-2i). The fact that induction in the test patches of White stimuli, as well as in simultaneous contrast and grating induction stimuli (Blakeslee & McCourt, 1997), is not homogeneous is of theoretical importance since virtually all models of brightness/lightness perception, except the ODOG and other closely related models relying on spatial filtering (Blakeslee & McCourt, 1999; Robinson, Hammond, & de Sa, 2007), predict and assume that induction is homogeneous within the test patches of these stimuli.

ACKNOWLEDGMENTS

This work was supported by grants from the National Institutes of Health NIH R01 EY014015 (BB) and NIH P20 GM103505 (MEM). The National Institute of General Medical Sciences and the National Eye Institute are components of the National Institutes of Health. The contents of this report are solely the responsibility of the authors and do not necessarily reflect the official views of the National Institutes of Health, National Institute of General Medical Sciences, or National Eye Institute. Commercial relationships: None.

REFERENCES

Adelson, E. H. (2000). Lightness perception and lightness illusions. In M. Gazzaniga (Ed.), *The new cognitive neurosciences* (2d ed.) (pp. 339–351). Cambridge, MA: MIT Press.

Anderson, B. (1997). A theory of illusory lightness and transparency in monocular and binocular images: The role of contour junctions. *Perception*, *26*, 419–453.

Anderson, B. L. (2001). Contrasting theories of White's illusion. *Perception*, *30*, 1499–1501.

Anderson, B. L. (2003). Perceptual organization and White's illusion. *Perception*, *32*, 269–284.

Anstis, S. (2006). White's effect in lightness, color, and motion. In M. Jenkins & L. Harris (Eds.), *Seeing spatial form* (pp. 91–100). New York: Oxford University Press.

Blakeslee, B., & McCourt, M. E. (1997). Similar mechanisms underlie simultaneous brightness contrast and grating induction. *Vision Research*, *37*, 2849–2869.

Blakeslee, B., & McCourt, M. E. (1999). A multiscale spatial filtering account of the White effect, simultaneous brightness contrast and grating induction. *Vision Research*, *39*, 4361–4377.

Blakeslee, B., & McCourt, M. E. (2003). A multiscale spatial filtering account of brightness phenomena. In L. Harris & M. Jenkins (Eds.), *Levels of perception* (pp. 47–72). New York: Springer.

Blakeslee, B., & McCourt, M. E. (2004). A unified theory of brightness contrast and assimilation incorporating oriented multiscale spatial filtering and contrast normalization. *Vision Research*, *44*, 2483–2503.

Blakeslee, B., & McCourt, M. E. (2005). A multiscale spatial filtering account of grating induction and remote brightness induction effects: Reply to Logvinenko. *Perception*, *34*, 793–802.

Blakeslee B., Pasieka, W., & McCourt, M. E. (2005). Oriented multiscale spatial filtering and contrast normalization: A parsimonious model of brightness induction in a continuum of stimuli including White, Howe and simultaneous brightness contrast. *Vision Research*, *45*, 607–615.

Bressan, P., & Kramer, P. (2008). Gating of remote effects on lightness. *Journal of Vision*, *8*(2), 1–8.

Carandini, M., Heeger, D. J., & Movshon, J. A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *Journal of Neuroscience*, *17*, 8621–8644.

DeValois, R. L., & DeValois, K. K. (1988). *Spatial vision*. New York: Oxford University Press.

Foley, J. M., & McCourt, M. E. (1985). Visual grating induction. *Journal of the Optical Society of America A*, *2*, 1220–1230.

Geisler, W. S., & Albrecht, D. G. (1995). Bayesian analysis of identification performance in monkey visual cortex: Nonlinear mechanisms and stimulus certainty. *Vision Research*, *35*, 2723–2730.

Gilbert, C. D., Das, A., Ito, M., Kapadia, M., & Westheimer, G. (1996). Spatial integration and cortical dynamics. *Proceedings of the National Academy of Sciences USA*, *93*, 615–622.

Gilchrist, A. L. (2006). *Seeing black and white*. New York: Oxford University Press.

Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., … Economou, E. (1999). An anchoring theory of lightness perception. *Psychological Review*, *106*, 795–834.

Helson, H. (1963). Studies of anomalous contrast and assimilation. *Journal of the Optical Society of America*, *53*, 179–184.

Howe, P. D. L. (2005). White's effect: Removing the junctions but preserving the strength of the illusion. *Perception*, *34*, 557–564.

Jameson, D. (1985). Opponent-colours theory in the light of physiological findings. In D. Ottoson & S. Zeki (Eds.), *Central and peripheral mechanisms of colour vision* (pp. 83–102). London: Macmillan.

Jameson, D., & Hurvich, L. M. (1989). Essay concerning color constancy. *Annual Review of Psychology*, *40*, 1–22.

Kingdom, F. A. A. (2011). Lightness, brightness and transparency: A quarter century of new ideas, captivating demonstrations and unrelenting controversy. *Vision Research*, *51*, 652–673.

Kingdom, F. A. A., & Moulden, B. (1988). Border effects on brightness: A review of findings, models and issues. *Spatial Vision*, *3*, 225–262.

Moulden, B., & Kingdom, F. A. A. (1989). White's effect: A dual mechanism. *Vision Research*, *29*, 1245–1259.

Ripamonti, K., & Gerbino, W. (2001). Classical and inverted White's effects. *Perception*, *30*, 467–488.

Robinson, A. E., Hammond, P. S., & de Sa, V. R. (2007). Explaining brightness illusions using spatial filtering and local response normalization. *Vision Research*, *47*, 1631–1644.

Rossi, A. F., & Paradiso, M. A. (1999). Neural correlates of perceived brightness in the retina, lateral geniculate nucleus, and striate cortex. *Journal of Neuroscience*, *19*, 6145–6156.

Rossi, A. F., Rittenhouse, C. D., & Paradiso, M. A. (1996). The representation of brightness in primary visual cortex. *Science*, *273*, 1104–1107.

Smith, V. C., Jin, P. Q., & Pokorny, J. (2001). The role of spatial frequency in color induction. *Vision Research*, *41*, 1007–1021.

Spehar, B., Clifford, C. W. G., & Agostini, T. (2002). Induction in variants of White's effect: Common or separate mechanisms? *Perception*, *31*, 189–196.

Spehar, B., Gilchrist, A. L., & Arend, L. E. (1995). The critical role of relative luminance relations in White's effect and grating induction. *Vision Research*, *35*, 2603–2614.

Taya, R., Ehrenstein, W., & Cavonius, C. (1995). Varying the strength of the Munker-White effect by stereoscopic viewing. *Perception*, *24*, 685–694.

Todorovic, D. (1997). Lightness and junctions. *Perception*, *26*, 379–395.

White, M. (1979). A new effect of pattern on perceived lightness. *Perception*, *8*, 413–416.

White, M. (1981). The effect of the nature of the surround on the perceived lightness of grey bars within square-wave test gratings. *Perception*, *10*, 215–230.

White, M., & White, T. (1985). Counterphase lightness induction. *Vision Research*, *25*, 1331–1335.

Yazdanbakhsh, A., Arabzadeh, E., Babadi, B., & Fazl, A. (2002). Munker-White-like illusions without T-junctions. *Perception*, *31*, 711–715.

Zaidi, Q., Spehar, B., & Shy, M. (1997). Induced effects of background and foregrounds in three-dimensional configurations: The role of T-junctions. *Perception*, *26*, 395–408.

Chapter 40
The Dungeon Illusion

Paola Bressan and Peter Kramer

COLORS DECEIVE US

We likely evolved color perception, in part, to better figure out whether or not food is good to eat (e.g., Regan et al., 2001; see also Kramer & Bressan, 2010a). In ordinary sunlight, for example, a ripe tomato reflects a certain portion of the light and looks red and edible; a rotten tomato reflects the light differently and looks brown and repulsive. It is these light-reflecting properties (*reflectance*) of the tomato—its redness or brownness[1]—that are important, not the fact that the tomato is illuminated by the sun; a red tomato under a blue lamp is just as edible. It so happens, however, that a red tomato absorbs blue light and does not reflect it. Consequently, a ripe tomato under a blue lamp does not look red but black, and rather unappetizing. This example illustrates the fundamental problem of color perception: it is impossible to separate a region's reflectance (color) from its illumination. Many animals take advantage of the confusion between color and illumination in their camouflage. These animals have darker backs than bellies (or the converse if they usually hang upside down, like some caterpillars). This *countershading*, whose remnants can be seen in dogs, offsets the body-shape revealing effects of light and shadow, rendering the animal less visible from a distance (Thayer, 1896).

In the niche of many species, the sun and the moon typically illuminate not just a few spots but the whole scene. Although there is no foolproof way to separate reflectance from illumination, it is this circumstance that makes a decent—although still fallible—color perception possible. The visual system's trick is to take the context into consideration (Bressan, 2006; Gilchrist et al., 1999). For example, the *luminance* of a sugar cube (i.e., the amount of light that the cube reflects) can be a billion times higher in the sun than in shadow, but even in the shadow the cube's *lightness* (its perceived achromatic color or shade of gray) remains white; a sugar cube is not perceived to change color every time we move it from the sun into shadow and back. The visual system's determination of the cube's lightness can be this stable thanks to a reliance not simply on the luminance of the cube but on the ratio between this luminance and that of the cube's surroundings. When illumination varies, this ratio (*contrast*) remains essentially the same. If, for instance, the luminance ratio between a sugar cube and the mahogany table on which it sits is 30:1 in shadow, it may become 30,000:1,000 in the sun, but that still amounts to the same proportion.

The reliance on luminance ratios, however, comes at a cost: contrast illusions. These arise whenever differences in the luminance of the surroundings are due to fluctuations not in illumination but in pigmentation. The familiar handbook example is a gray square that looks lighter on a black than on a white background, but contrast illusions abound (see Fig. III.40-1). They are so ubiquitous that lightness illusions that seem to *reverse* them are especially interesting, and the dungeon illusion is one of these.

THE DUNGEON ILLUSION AND ITS VARIATIONS

The Dungeon Illusion

The dungeon illusion is made of two displays, each containing identical gray disks that are surrounded by white disks on a black background and by black disks on a white background, as in Figure III.40-2a–b. The gray disks look lighter on the white background (Bressan, 2001, 2006, 2007a; Bressan & Kramer, 2008), which is an effect opposite to that of contrast. The illusion is so named because the original version had squares in place of disks (Bressan, 2001); the grid created by the space between the squares made the figure look like a pair of prison windows.

Double-Decrement Dungeon Illusion

In this variant, the central disks (which we call *targets*) are dark gray and are surrounded by white disks on a gray background and by gray disks on a white background, as in Figure III.40-2c–d. The targets, being dimmer than both the contextual disks and the backgrounds, are called *double decrements*. Here the dungeon illusion reverses: as in a regular contrast effect, the targets look lighter on the gray background than on the white one.

Interestingly, the double-decrement dungeon illusion is larger if the environment is black rather than white (Bressan & Kramer, 2008). This goes to show that a well-lit picture, designed to convey a particular lightness effect, may not produce such an effect if viewed on white paper in



Figure III.40-1. Left panel: The two disks under the chair are the same color as those on the mirror, and the brightest of the two disks under the chair and the dimmest of the two disks in front of the window are the same color too. The right panel is a negative of the left one except for the disks, which are exactly the same as those on the left. The color illusions contained in the left panel do not emerge in the right one. It is the context that makes the difference. (Adapted from Bressan, 2007b, p. 91.)

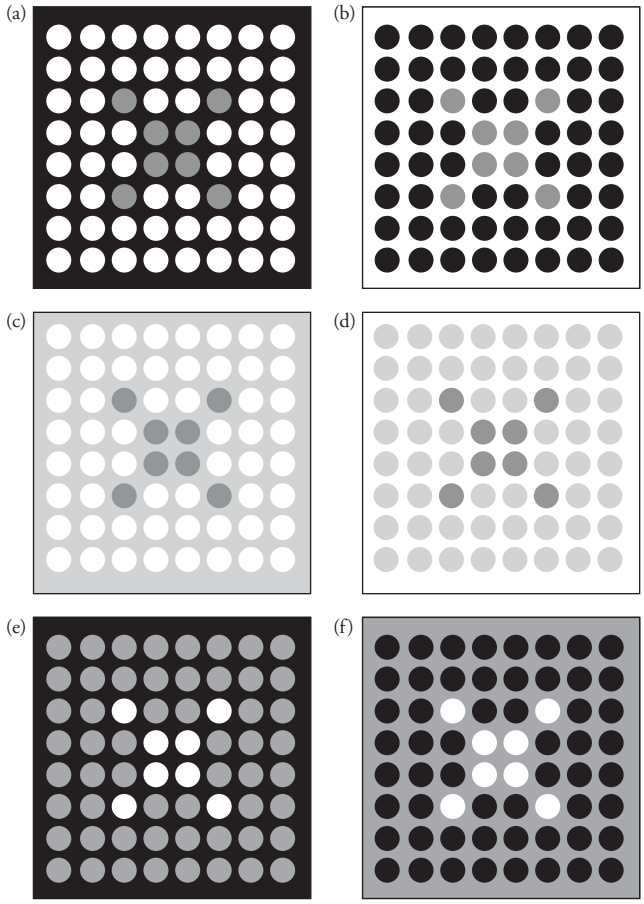


Figure III.40-2. (a) and (b): The dungeon illusion. The eight gray disks (called targets) in (a) and (b) have the same luminance but look darker in A than in B: an effect opposite to that of contrast. (c) and (d): The double-decrement dungeon illusion (best seen in a completely black environment). The dark gray targets in (c) and (d) have the same luminance but look darker in (d) than in (c): an effect opposite to the dungeon illusion and in the same direction as that of contrast. (e) and (f): The double-increment dungeon illusion (best seen in a completely white environment): the light gray targets in (c) and (d) have the same luminance but look lighter in (e) than in (f): again, an effect opposite to the dungeon illusion and in the same direction as that of contrast. (Adapted from Bressan & Kramer, 2008, p. 3.)

a bright room but may suddenly work if viewed on black paper in a dark room. An extreme, stunning example is the Gelb (1929) effect: in a dark room, a well-lit black disk does not look black but white. As soon as the lights are turned on, or a large piece of white paper is placed behind the disk, the disk immediately appears to turn black. If the white paper is removed, however, then the black disk will once again look white. In the same way, the moon can look white and shiny in the earth's night sky and yet dark gray under an astronaut's boots (see also Bressan, 2005). Curiously, and—as explained later—importantly, variations in the environment's brightness affect the figure with the gray background more than the other (e.g., Fig. III.40-2c more than Fig. III.40-2d).

Double-Increment Dungeon Illusion

In this variant, the targets are light gray and are surrounded by gray disks on a black background and black disks on a gray background, as in Figure III.40-2e–f. The targets, being brighter than both the contextual disks and the backgrounds, are called *double increments*. Again, the dungeon illusion reverses: the targets look lighter on the black background than on the gray one. The double-increment illusion emerges if the environment is white but not if it is black (Bressan & Kramer, 2008). Again, curiously and importantly, variations in the environment's brightness affect the figure with the gray background more than the other (e.g., Fig. III.40-2f more than Fig. III.40-2e).

Paradoxical Dungeon Illusion

If Figure III.40-2d is placed on a larger black surface and the rest of the room is dark, then the dark-gray targets look *darker* in the presence than in the absence of the gray contextual disks. If Figure III.40-2e is placed on a larger white surface, then the light-gray targets look *lighter* in the presence than in the absence of the gray contextual disks (Kramer & Bressan, 2010b). This effect is paradoxical because, contrary to any contrast-based expectation, decreasing the overall brightness of the surround darkens rather than lightens objects, and increasing the overall brightness

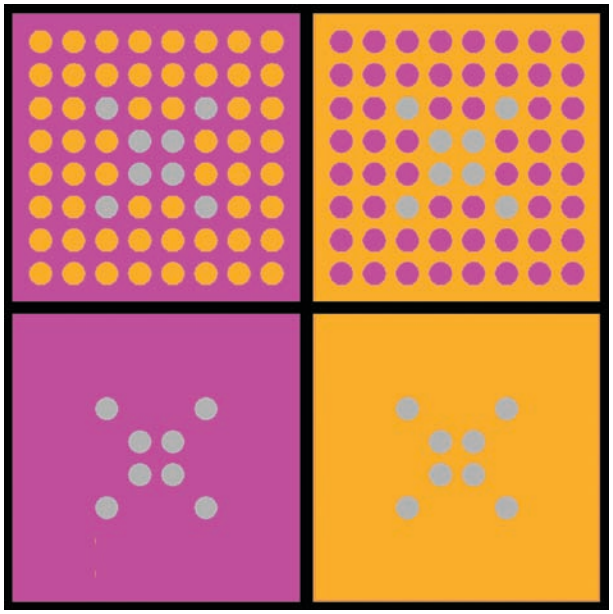


Figure III.40-3. In each of the four panels, the central targets are gray. Yet the targets look bluish in the top left panel and greenish in the top right one: a color variant of the dungeon illusion. However, the targets look greenish in the bottom left panel and bluish in the bottom right one, showing that the illusion above is due to the contextual disks. The effects become stronger with longer viewing. Small elements tend to become more similar to their background (assimilation), which adds to the effect: consequently, the smaller the figure is, or the farther away one is relative to it, the larger the illusion. (Adapted from Bressan & Kramer, 2008, p. 7.)

of the surround lightens rather than darkens them. This effect differs not only from contrast illusions but also from reverse-contrast ones. Never had it been shown that regions can lighten when the average luminance of the entire scene around them is increased rather than decreased.

Chromatic Dungeon Illusion

If white and black are replaced with orange and magenta, as in the top panels of Figure III.40-3, the gray disks look respectively bluish (on the magenta background) and greenish (on the orange background). If all surrounding disks are removed, however, as in the bottom panels of Figure III.40-3, the gray disks become respectively greenish on the magenta background and bluish on the orange one—as in a normal color contrast effect (Bressan & Kramer, 2008).

THE DUNGEON ILLUSION'S IMPORTANCE

The dungeon illusion defied theories of lightness that were popular at the time it was presented. Several current models, including Gilchrist et al.'s (1999) anchoring theory, are challenged by the dungeon illusion's variants, although other more recent models (e.g., edge-integration models: Vladusich, Lucassen, & Cornelissen, 2006) may perhaps be able to deal with them. The double-anchoring model of lightness (Bressan, 2006) is consistent with both the dungeon illusion and its variants. More important, it was this model that led to the discovery of both the existence of the variants and the surprising effects that the surrounding environment has on them.

The Double-Anchoring Model of Lightness

Contrast illusions are not the only drawback of a reliance on contrasts. Another disadvantage is that contrast information tells our visual system only how much one object is lighter than another, not what the object's color is. If one object is twice as light as another, for example, then it could be white and the other gray, or it could be gray and the other black, and so on; the number of pairs of luminances whose ratio is 2:1 is virtually infinite. What our visual system needs is some form of *anchoring*, so as to establish at least one point of contact between the luminance scale (spanning from the lowest to the highest luminance in a scene) and the lightness scale (spanning from the color black to the color white).

In Bressan's (2006) double-anchoring model of lightness, which evolved out of Gilchrist et al.'s (1999) (single-) anchoring model, the lightness of a region depends on its grouping with other regions in the scene and on two types of anchoring: highest-luminance anchoring and surround anchoring. Grouping between different regions expresses how likely they are to share the same illumination (and should not be confused with classic perceptual grouping, which expresses how likely regions are to be part of the same object; see Bressan's, 2007a, response to criticism by Howe, Sagreiya, Curtis, Zheng, & Livingston, 2007). The more similarly illuminated two regions are, the more probable it is that luminance differences between them are due to differences in their reflectances (colors) rather than to differences in illumination. Hence, comparing a target region with regions that are similarly, as opposed to differently, illuminated produces a more reliable estimate of its reflectance. This is why, in the model, the better two regions group, the more they influence each other's lightness.

Let us take an example. White regions, such as snow, typically reflect 90% of the light that hits them and black regions, such as asphalt, less than 3%. Yet the luminances of a patch of snow and a patch of asphalt in the sun are much more similar than the luminances of a patch of snow in the sun and an identical patch of snow in shadow. An important grouping principle is therefore *luminance similarity*: the more similar in luminance two regions are, the better they group. Other important grouping principles are luminance polarity, proximity, and coplanarity.

Consider now a target region and some other region that groups well with it (in the model called *surround*). For this and all other pairs that group well, the model performs two kinds of anchoring. *Surround anchoring* assigns to the target a provisional lightness that is equal to the luminance ratio between the target and the surround. *Highest-luminance anchoring*, instead, assigns to the target a provisional lightness that is equal to the luminance ratio between the target and the region with the highest luminance (either the surround or the target itself). The target thus receives two provisional lightnesses in each target–surround pair, one via surround anchoring and one via highest-luminance anchoring. The larger and brighter the

surround, the more weight is given to surround anchoring rather than highest-luminance anchoring. And the stronger the grouping between the target and a given surround, the more weight is given to the anchoring in this, rather than in another, target–surround pair. The target's final lightness is the weighted average of all its provisional lightnesses, computed across all target–surround pairs that group well.

Because a weighted average of luminance ratios is still just a luminance ratio and not a lightness, anchoring involves a final step: the luminance ratio is multiplied by the luminance of the color white. Then the determination of the target's lightness is complete.

Predicting the Double-Increment Contrast Illusion

The effects of double anchoring are most clearly seen in the double-increment contrast illusion (Bressan & Actis-Grosso, 2001): a gray region looks lighter on a black than on a dark-gray background. Although the illusion is simple, highest-luminance anchoring alone (as implemented in the single-anchoring model by Gilchrist et al., 1999) cannot explain it. Highest-luminance anchoring alone erroneously predicts that the gray region will always be perceived as white, because it is the highest luminance regardless of whether the background is black or dark gray. In surround anchoring, however, the target is computed as lighter on black than on dark gray, because the luminance ratio between the target and the background is larger when the background is black than when it is dark gray. A particularly convincing double-increment contrast illusion is the Christmas wall-of-blocks (Bressan, 2006; see Fig. III.40-4).

Predicting the Dungeon Illusion's Variants

In the dungeon illusion, background and contextual disks are either, respectively, brighter and dimmer or dimmer and brighter than the targets, and thus they have opposite effects on them. The dungeon illusion emerges when the targets group better with the contextual disks than with the background and reverses when the targets group better with the background than with the contextual disks (see Bressan, 2007a). Take Figure III.40-2a–b. The targets and the contextual disks are both brighter (Fig. III.40-2a) or both dimmer (Fig. III.40-2b) than the background. They have the same *luminance polarity* and group well (Masin, 2003). Indeed, the contextual disks group so well with the targets that their effect outweighs that of the background, creating the dungeon illusion.

Things are quite different in the double-decrement and double-increment pairs. In the pair member with the gray background, either the targets are dimmer and the contextual disks brighter (Fig. III.40-2c) or the targets are brighter and the contextual disks dimmer (Fig. III.40-2f) than the background. They have opposite luminance polarities and group poorly (Masin, 2003); this weakens the influence of the contextual disks on the targets. Additionally, in the pair member with the white background the contextual disks are also brighter than the targets (Fig. III.40-2d), and in the pair member with the black background the contextual disks are also dimmer than the targets (Fig. III.40-2e). Hence, in these figures the effects of the contextual disks are not even opposite to the background's but go in the same direction (target darkening in Fig. III.40-2d and target lightening in Fig. III.40-2e). In the double-decrement and double-increment pairs, therefore, the contextual disks should produce either a weaker dungeon illusion or a reversed one (i.e., a regular contrast effect), which is indeed the case.

Predicting the Paradoxical Dungeon Illusion

The effect of weight assignment to luminance ratios is most clearly seen in the paradoxical dungeon illusion. Here we have a *local* background (white in Fig. III.40-2d, black in Fig. III.40-2e) and a *remote* background, that is not shown in Figure III.40-2 but that surrounds the local background and has the opposite color (black in Fig. III.40-2d, white in Fig. III.40-2e).

Consider Figure III.40-2e (a similar argument holds for Fig. III.40-2d). The paradoxical dungeon illusion emerges when the remote background is white and large enough to affect the targets more than the local background does. The local background is black and lightens the targets; the remote background is white and darkens them. Because the effect of the remote background is larger than that of the local background, the net effect of the two backgrounds combined is a darkening of the targets. If present, the contextual disks are dimmer than the targets, and hence they do lighten them but of course less than the parts of the black background they replace. Thus the targets should look darker with the contextual disks than without them. Surprisingly, they do not; they look lighter.

The double-anchoring theory explains this phenomenon as follows. The luminance ratios between (a) target and local background, (b) target and remote background, and (c) target and contextual disks, if present, are each assigned a relative weight that expresses their relative importance in determining the target's lightness. The relative

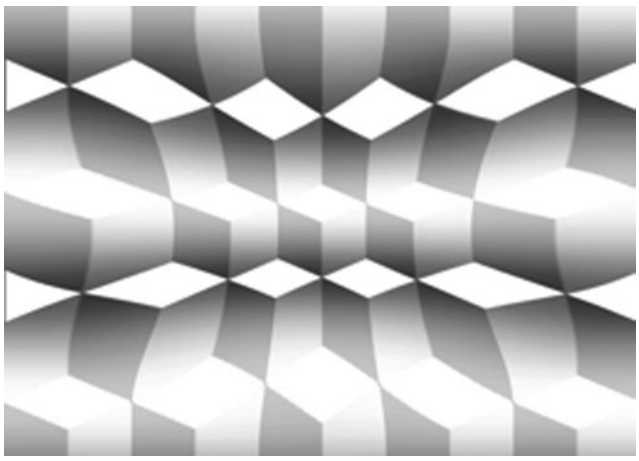


Figure III.40-4. The Christmas wall of blocks. All diamonds are the same color, but those in the first and third rows look lighter than those in the second and fourth. Critically, all diamonds are the highest luminance in the scene. (Adapted from Bressan, 2006, p. 528.)

weights sum up to 1 (by definition, they sum up to a constant). Consequently, when the contextual disks are present, and some weight is assigned to their luminance ratio with the targets, then some weight must be taken away from the luminance ratios of the backgrounds with the target. Thus the net darkening effect of the backgrounds receives less weight when the contextual disks are present than when they are absent. If the remote background is large and bright enough, then the paradoxical result is that the targets look lighter in the presence than in the absence of the contextual disks.

Predicting the Effects of the Environment on the Dungeon Illusions

In the model, the weights given to luminance ratios depend on grouping. The effect of grouping on weight assignments is most clearly seen in the larger influence that the environment has on Figure III.40-2c than on Figure III.40-2d and on Figure III.40-2f than on Figure III.40-2e. That is, the remote background influences the figures with the gray backgrounds more than the others.

The reason for this becomes clear when we recall that, in the paradoxical dungeon illusion, targets are more affected by the remote background when there are no contextual disks. Notice that in the figures with the gray backgrounds the grouping between targets and contextual disks is weaker, because they have opposite luminance polarities. If they did not group with each other at all, one could just as well remove the contextual disks. As the paradoxical contrast illusion shows, it is in this case that the environment has the largest effect on the targets. So it makes sense that the remote background affects target lightness more in Figures III.40-2c and III.40-2f than in Figures III.40-2d and III.40-2e.

Why Should Double Anchoring Have Evolved?

We may wonder why, to perceive colors, we should have taken the trouble of evolving *two* types of anchoring rather than just one. The answer may have to do with some aspects of nature that are crucial to our survival: light sources and highly reflective surfaces (Bressan, 2006). The most important of these are, perhaps, fire and water. Water is transparent, but its surface acts as a natural mirror that is more reflective than almost anything else in the natural world. We can recognize the sparkling surface of water from a long distance, and the benefits of this ability are obvious. Fire does not reflect light but emits it; recognizing this property helps us to identify it as a heat source. Fire is both uniquely useful and dangerous, and there is little doubt that it has been, and still is, exceptionally important to our survival. To a visual system based only on surround anchoring—in an achromatic but otherwise natural world—all regions that are brighter than their surrounds would look whiter than white (luminous); to a visual system based only on highest-luminance anchoring, none ever would. Sparkling waters and raging fires look luminous thanks to surround anchoring, but it is via highest-luminance anchoring that they can be discriminated from all other bright surfaces and pop out.

NOTE

1. "Redness" and "brownness" refer to colors; here we define "color" as the spectrum of reflected or emitted light (for alternative definitions and their presumed implications, see Byrne & Hilbert, 2003).

REFERENCES

Bressan, P. (2001). Explaining lightness illusions. *Perception*, *30*, 1031–1046.

Bressan, P. (2005). The dark shade of the moon. *Clinical and Experimental Ophthalmology*, *33*, 574.

Bressan, P. (2006). The place of white in a world of grays: A double-anchoring theory of lightness perception. *Psychological Review*, *113*, 526–553.

Bressan, P. (2007a). Dungeons, gratings, and black rooms: A defense of double-anchoring theory and a reply to Howe et al. (2007). *Psychological Review*, *114*, 1111–1115.

Bressan, P. (2007b). *Il colore della luna: Come vediamo e perché*. Roma-Bari: Laterza.

Bressan, P., & Actis-Grosso, R. (2001). Simultaneous lightness contrast with double increments. *Perception*, *30*, 889–897.

Bressan, P., & Kramer, P. (2008). Gating of remote effects on lightness. *Journal of Vision*, *8*(2), 1–8.

Byrne, A., & Hilbert, D. R. (2003). Color realism and color science. *Behavioral and Brain Sciences*, *26*, 3–64.

Gelb, A. (1929). Die "Farbenkonstanz" der Sehdinge. In W. A. von Bethe (Ed.), *Handbuch der normal und pathologische psychologie* (Vol. 12, pp. 594–678). Berlin: Springer.

Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., & Li, X., . . . Economou, E. (1999). An anchoring theory of lightness perception. *Psychological Review*, *106*, 795–834.

Howe, P. D., Sagreiya, H., Curtis, D. L., Zheng, C., & Livingstone, M. S. (2007). The double-anchoring theory of lightness perception: A comment on Bressan (2006). *Psychological Review*, *114*, 1105–1109; discussion 1111–1115.

Kramer, P., & Bressan, P. (2010a). Ignoring color in transparency perception. *Rivista di Estetica*, *43*, 147–159.

Kramer, P., & Bressan, P. (2010b). Paradoxical lightness contrast. *Vision Research*, *50*, 144–148.

Masin, S. C. (2003). Effects of algebraic and absolute luminance differences on achromatic surface grouping. *Perception*, *32*, 615–620.

Regan, B. C., Julliot, C., Simmen, B., Viénot, F., Charles-Dominique, P., & Mollon, J. D. (2001). Fruits, foliage and the evolution of primate color vision. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, *356*, 229–283.

Thayer, A. H. (1896). The law which underlies protective coloration. *Auk*, *13*, 124–129.

Vladusich, T., Lucassen, M. P., & Cornelissen, F. W. (2006). Edge integration and the perception of brightness and darkness. *Journal of Vision*, *6*, 1126–1147.

Chapter 41
The Contrast Illusion

Charles Chubb, Joshua A. Solomon, and George Sperling

INTRODUCTION

The contrast contrast illusion is shown in its basic form in Figure III.41-1a (Chubb, Sperling, & Solomon, 1989). The two central disks of texture on the left and right are physically identical. However, to most observers, the left one (appearing on the mean gray background) appears higher in contrast than the right one (appearing on the high-contrast texture background).

The contrast contrast illusion is analogous to the classical simultaneous contrast illusion shown in the small image inset in the left side of Figure III.41-1a. In this figure, the two disks are physically identical; however, the one on the left (seen against a black background) appears brighter than the one on the right (seen against a white background). This illusion illustrates that the brightness of a homogeneous test patch depends dramatically on the context in which the test patch is situated. The contrast contrast illusion shows that the same is true of the apparent contrast of a textured test patch: a test patch filled with medium-contrast texture appears higher in contrast when viewed against a homogeneous gray field (which we can think of as a background texture of contrast zero) than it does against a background field filled with high-contrast texture.

A standard account of the simultaneous contrast illusion (first given by Mach, 1866) proposes that the neurons registering the intensity of light impinging on nearby points of the retina laterally inhibit each other. Thus, for example, the neurons stimulated by the light from the test patch on the right side of the small image inset in Figure III.41-1a are inhibited by the more strongly activated neurons stimulated by the white background. By contrast, the neurons stimulated by the light from the test patch on the left side of the inset in Figure III.41-1a receive no such inhibition from the neurons stimulated by the black background; as a result, the patch on the left appears brighter than the patch on the right. Most of the computational models offered to account for the contrast contrast effect (e.g., Chubb, Sperling, Solomon, 1989; Heeger, 1994; Singer & D'Zmura, 1995) are analogous to Mach's (1866) lateral inhibition model of brightness perception.

SENSING TEXTURE CONTRAST

To grasp the spirit of this model, however, it is necessary first to understand how the visual system senses the contrast of a texture. To measure the intensity of a light, one might use a simple meter that counts quanta impinging on a region of the retina per unit time; indeed, retinal receptors (rods and cones) do something like this. This strategy will not work to measure the contrast of a texture for the simple reason that one can increase the contrast of a texture without changing the average light intensity of the texture. As the bright points of the texture are increased in intensity, the dark points are decreased so that the average intensity does not change. To gauge the contrast of a texture, one must extract from the image of the texture a statistic that reflects the strength with which intensities *vary* inside the texture.

Many different neurons in the visual processing stream provide potential access to information about texture contrast. At the level of the retina, on-center retinal ganglion neurons (on-cells) and off-center retinal ganglion neurons (off-cells) will both capture information about texture contrast in the sense that each population of neurons will fire more strongly in response to high-contrast than to low-contrast texture. The reason is simple, as we illustrate first in the case of on-cells.

As shown by the magenta lines in Figure III.41-2a, an on-cell has a receptive field with a small excitatory central region (i.e., light impinging on this region tends to increase the firing rate of the neuron) and a larger, doughnut-shaped, inhibitory surround (i.e., light impinging on this region tends to decrease the firing rate of the neuron), and the on-cell's response will be roughly proportional to the difference between the average intensity of the light impinging on its center versus its surround *provided* that this difference is positive (i.e., provided the light impinging on its center is more intense than on its surround); if the difference is negative, the on-cell will not fire at all. Because the light inside the central magenta circle in Figure III.41-2a is brighter than the light falling in the doughnut-shaped surround, an on-cell whose receptive field matched the one illustrated here would fire strongly.

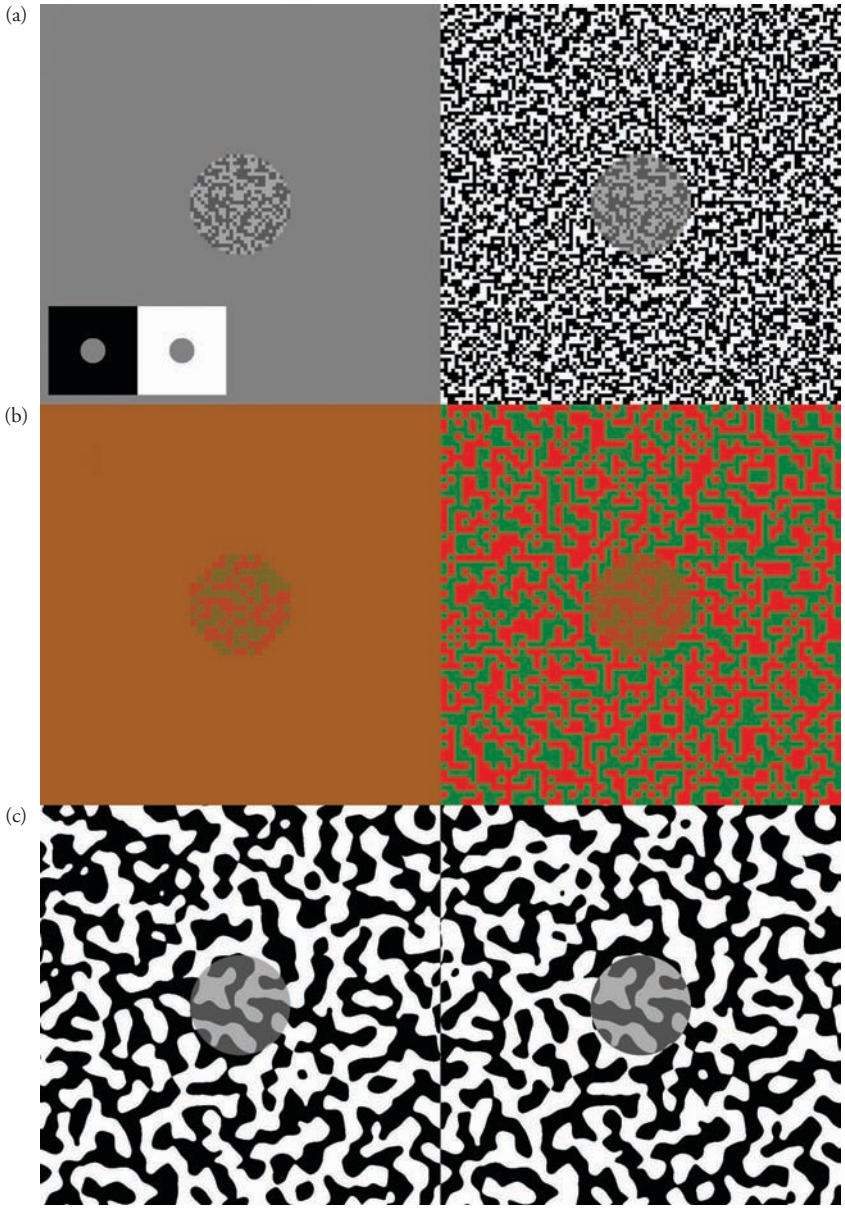


Figure III.41-1. Three variants of the contrast contrast illusion. In each case, the test disks appearing on each of the left and right sides are physically identical; however, to most observers the disks on the left appear higher in contrast than the disks on the right. (a) A medium contrast disk of achromatic texture appears higher in contrast when it appears against a uniform gray background than it does when it appears against a high-contrast texture background; this display is analogous to the simultaneous contrast display inset into the left side. (b) A medium contrast disk of equiluminant, red-green texture appears higher in contrast when it appears against a uniform brown background than it does when it appears against a high-contrast texture background. (c) If the disk can be interpreted as a translucent (right panel), its texture appears lower in contrast than it does if boundary cues preclude such an interpretation (left panel).

An off-cell works like an on-cell except that its surround is excitatory and its center is inhibitory. As this implies, the off-cell will respond above zero only if the light hitting its central region is lower in intensity than the light hitting its surround. Thus, for example, an off-cell with the receptive field indicated by the green lines in Figure III.41-2a would give a strong positive response because the light hitting its center is lower in intensity than the light hitting its doughnut-shaped surround.

Why does the visual system use both on-cells and off-cells? It seems it would make more sense to use just one type of neuron whose firing rate was proportional to the difference between the average light intensity in the central, excitatory region minus the average light intensity in the doughnut-shaped surround. Although this strategy seems natural, it implies that whenever the light in the central region is lower in intensity than the light in the surround, the neuron must be able to lower its firing rate below zero, which in turn implies that the neuron must maintain a resting firing rate (the firing rate that signals equal light levels in center and surround) that is greater

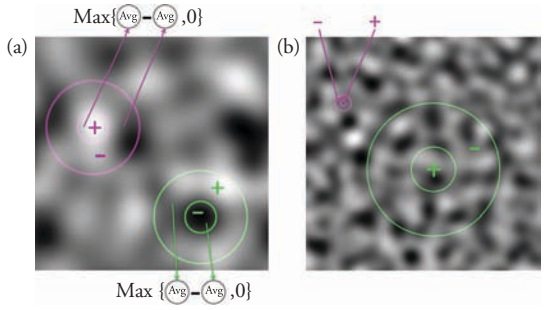


Figure III.41-2. How on- and off-center neurons respond to texture. (a) The magenta lines show the computation performed by a (strongly activated) on-center neuron, and the green lines show the computation performed by a (strongly activated) off-center neuron. (b) If the receptive field of an on-center neuron is too small (magenta) or too large (green) relative to the granularity of the texture, then the neuron will not be strongly activated.

than zero. The problem is that it takes resources to produce action potentials; thus it would be very wasteful for the visual system to use large arrays of neurons whose resting firing rate were elevated substantially above zero. Instead, the visual system avoids this waste by using separate channels (on-cells and off-cells), each with a resting firing rate of zero, for conveying information about positive versus negative deviations of light intensity away from the local average light level (Hubel, 1960; Kuffler, 1953).

Each population of neurons, the on-cells and the off-cells, provides information about texture contrast. This is illustrated in Figure III.41-3. Panel (a) shows a version of the contrast contrast illusion. Note that all of the main components of this figure (the homogeneous gray background on the left, the high-contrast texture background on the right, and the test disk of medium-contrast texture reproduced in the middle of each side of panel [a]) have equal average light intensity; blacks and whites vary symmetrically around mean gray. Panel (b) shows (downscaled in size) the pattern of output that results from applying a dense array of on-cells to the stimulus in panel (a). Each pixel value in this image represents the firing rate of an on-cell whose receptive field is centered at the corresponding pixel of panel (a). A firing rate of zero is coded by the same gray that is used for the homogeneous background in panel (a), and on-cell responses greater than zero are coded by grays of increasing brightness. Panel (c) shows the analogous pattern of output from a dense array of off-cells.

Note that regions of increasing texture contrast in panel (a) result in regions of increasing average activation in each of panels (b) and (c). The homogeneous gray background on the left side of panel (a) produces a homogeneous region of response zero from each of the on-cell and off-cell neural arrays. The test disk reproduced on both sides of panel (a) produces a smattering of medium-level activation in each of the on-cell and off-cell neural arrays, and the high-contrast texture background on the right of panel (a) produces a smattering of higher activation in each of the two arrays.

Note, however, that if any given pixel in the "neural image" (Robson, 1980) produced by the on-cells is non-zero, then the corresponding pixel in the off-cell neural image must be zero and vice versa. As this observation suggests, we can obtain a neural image whose local average level of activation signals contrast more continuously over space by adding the images in panels (b) and (c) together. The resulting image is shown in Figure III.41-3d. Many neurons in primary visual cortex (e.g., complex cells; Hubel & Wiesel, 1962) operate by additively combining on-cell and off-cell responses.

Suppose there exists in human visual cortex an array of neurons whose response to the contrast contrast stimulus shown in Figure III.41-3a is the neural image shown

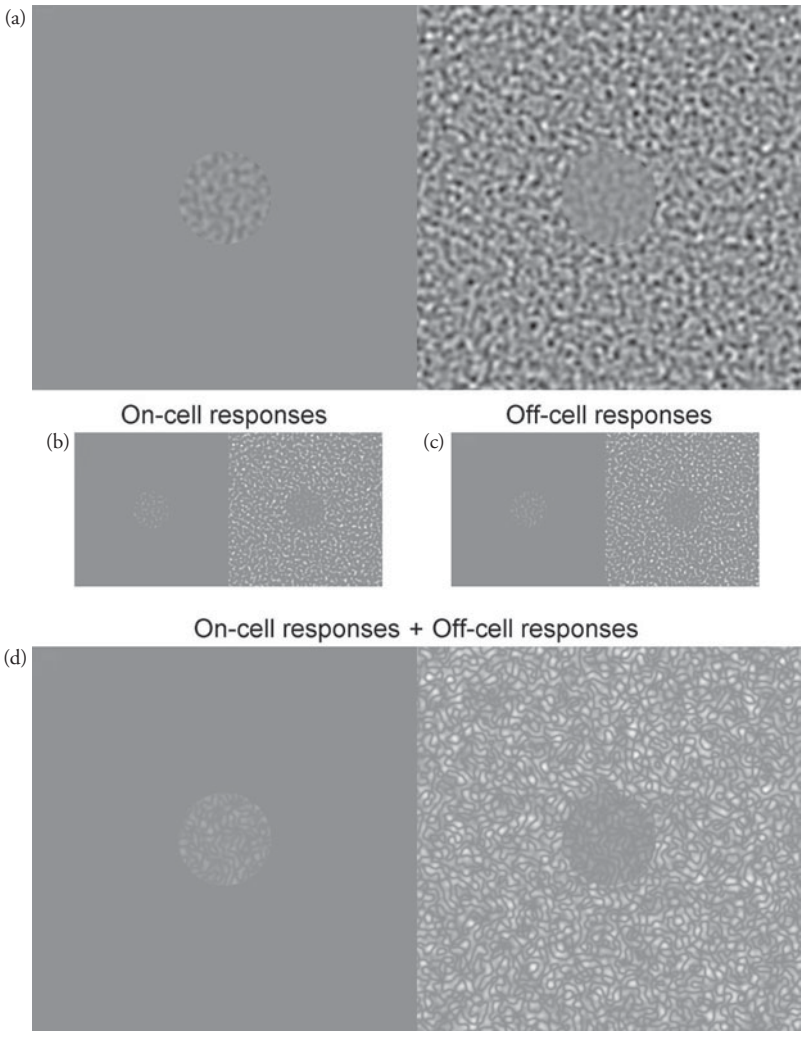


Figure III.41-3. How on- and off-center neurons could convert the contrast contrast display into a neural image akin to the simultaneous contrast display. (a) A version of the contrast contrast display. (b) Responses of on-center neurons to the image shown in (a). (c) Responses of off-center neurons to the image shown in (a). (d) The sum of the images in (b) and (c)—note the similarity between this image and the simultaneous contrast display inset in the left side of Figure III.41-1a.

in Figure III.41-3d. In itself, the existence of such a neural array does not explain the contrast contrast illusion. We must further hypothesize that the response of a given neuron in this array operates to suppress the responses of on-cells and off-cells whose receptive fields are in the same retinal neighborhood as the receptive field of the neuron. If this is the case, then the neurons monitoring the region of the visual field filled with the homogeneous gray background on the left in Figure III.41-3a will (as we see in Fig. III.41-3d) remain silent; by contrast, the neurons stimulated by the high-contrast texture background on the right will fire strongly. This means that the on-cells and off-cells sensing the test disk on the right will be inhibited much more strongly (from the highly activated neurons monitoring the nearby, high-contrast background texture) than the neurons sensing the test disk on the left.

THE IMPORTANCE OF TEXTURE GRANULARITY

In generating the images in Figures III.41-3B, C, and D, we used on-cells and off-cells whose receptive fields are matched in scale to the granularity of the texture being viewed. The receptive field centers of these on- and off-cells are roughly equal in size to the blobs in the texture (as is evident in Fig. III.41-2a). Figure III.41-2b shows what happens if this were not true. If the granularity of the texture were too fine relative to receptive field scale (as is the case for the receptive field indicated in green), then regardless of how the receptive field happened to align with the texture, its response would be very near zero because bright and dark variations within the receptive field center as well as in the receptive field surround would tend to average out to the same value. On the other hand, if the texture were too coarse relative to the receptive field scale (as is the case for the on-center receptive field indicated in magenta), then regardless of how the receptive field happened to align with the texture, its center and surround would be stimulated by light of approximately uniform, equal intensity, implying that its response would again be very near zero.

If the lateral inhibition hypothesized to underlie the contrast contrast illusion occurred only within neurons with receptive fields of a fixed spatial scale, then the suppression of the apparent contrast of the test patch in a contrast contrast display should be weakened when the granularity of the surrounding texture did not match the granularity of the texture in the test patch. In Figure III.41-4 the test patch is identical in the left, center, and right displays. However, the three background textures differ in granularity: the left background texture is twice as fine as the test patch; the center background is equal in granularity to the test patch; and the right background texture is twice as coarse as the test patch. Chubb, Sperling, and Solomon (1989) showed that, indeed, the apparent contrast of the test patch is suppressed more strongly (by roughly a factor of 2) when the granularity of the background matches that of the test patch than in either of the two cases shown on the left and right sides of Figure III.41-4 in which the granularity of target and background differ by a factor of 2.

INDIVIDUAL DIFFERENCES

Of course it can be difficult—if not impossible—to measure apparent contrast (or apparent anything, for that matter) without contamination from cognitive biases and/or expectation effects. What we can say for sure is that there are several results consistent with substantial individual differences in how the contrast contrast illusion seems to be experienced. Cannon and Fullenkamp (1993) reported one participant for whom the basic contrast contrast effect was dramatically reversed. For this participant, embedding a low-contrast patch of texture in a high-contrast background actually made the contrast of the texture patch appear higher than it did when it was viewed against a homogeneous gray background. For a second participant, embedding a small target texture patch in a larger surrounding texture of equal or lower contrast enhanced the contrast of the texture patch compared to the no-surround condition.

Nonetheless, surround-induced enhancement of apparent contrast seems to be relatively rare. Bosten and Mollon (2010) measured the contrast contrast illusion (and nine other illusions) in a pool of 100 normally sighted observers. For some of these observers, a textured surround only slightly suppressed the apparent contrast of the texture patch it contained; for others, it strongly suppressed the

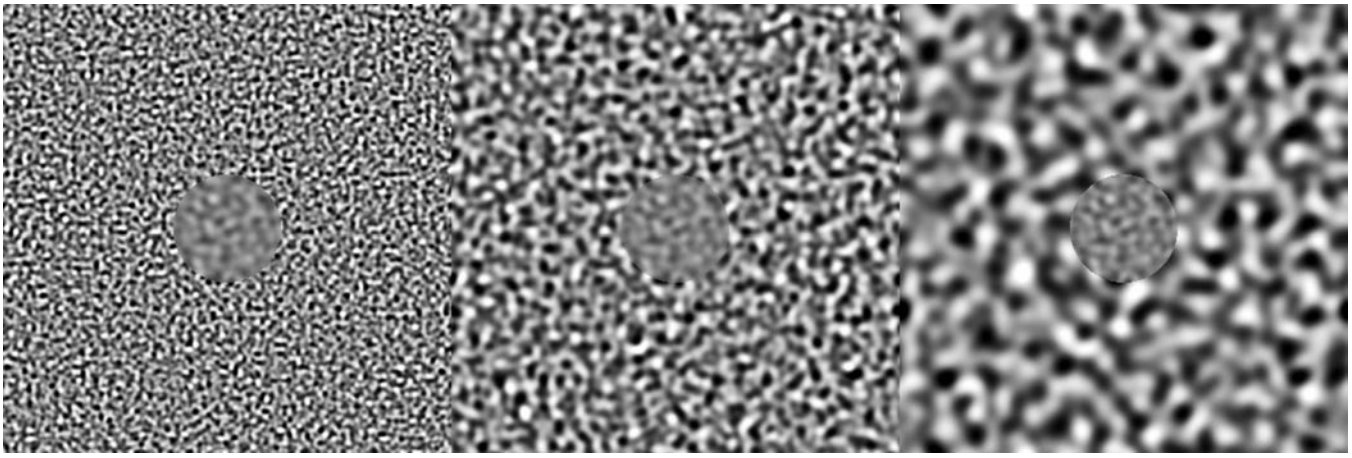


Figure III.41-4. How the granularity of the background texture influences the illusion. The background texture on the left (right) is finer (coarser) than texture in the test patch. Most observers judge the contrast of the right and left test patches to be higher than that of the central test patch, which is matched in granularity to its background texture.

apparent contrast of the target; but for none did it actually enhance the target's apparent contrast. Across observers, a correlation was found between the strength of this suppression and that caused by chromatic (red/green) contrast, but there was no significant correlation between the basic (luminance) contrast contrast effect and any of the other eight illusions studied.

Dakin, Carlin, and Helmsley (2005) reported results from 11 forensic inpatients for whom the basic contrast contrast effect was weak or absent. All of the observers suffered from schizophrenia. Yang et al. (2012) recently replicated this basic result in 30 outpatients. Dakin et al. hypothesized that abnormally low lateral inhibition may be related to both the reduced contrast contrast illusion and schizophrenia. However, two later studies of 132 outpatients with schizophrenia or schizoaffective disorder found neither a significant difference between the strengths with which normal observers and these patients experience the contrast contrast effect (Barch et al., 2012) nor any relationship between the severity of schizophrenic symptoms and the strength of the contrast contrast illusion (Gold et al., 2012).

IS THE CONTRAST CONTRAST ILLUSION SELECTIVE FOR CONTRAST POLARITY?

In our previous discussion (see Fig. III.41-2) we suggested that the contrast contrast illusion might be mediated by neurons that sum the responses of on-cells and off-cells together to create a neural image whose response strength at a given point reflects the contrast of the texture in the neighborhood of that point. Under this hypothesis, the neural substrate that gives rise to the illusion has discarded all information about whether the contrast registered at a particular point was signaled by an on-cell or an off-cell (i.e., whether luminance deviated negatively or positively from the local average). Is this true?

Although the textures we used in the stimulus displays in Figures III.41-1a, III.41-2a, and III.41-4 drive on-cells and off-cells with equal effectiveness, it is possible to create textures that activate one of these two classes of neurons more strongly than the other. This is exactly what Sato, Motoyoshi, and Sato (2012) did. Their stimulus textures are shown in Figure III.41-5. Each of these textures contains either dark vertical elements or bright vertical elements on a gray background. The textures with the bright bars are designed to stimulate on-cells more strongly than off-cells and vice versa for the textures with the dark bars. Notice that the backgrounds differ in their brightnesses; the brighter the vertical elements, the darker the background. This is to ensure that all of the textures have the same the average light intensity.

Compare the rightmost upper and lower stimuli. Most likely, as the participants Sato et al. (2012) used in their experiments experienced, the central patch of texture in the upper rightmost stimulus will appear lower in contrast than the (equal contrast) patch of texture in the lower rightmost stimulus. Sato et al. conclude that on-cells are laterally inhibited by on-cells more effectively than they are by off-cells, contradicting the hypothesis (see Fig. III.41-5) that the site of the inhibition is a neural substrate in which on-cell and off-cell responses have been additively combined.

Notice that the difference in the apparent contrast of the central texture patches in the stimuli on the far left of Figure III.41-5 is less dramatic than it is in the stimuli on the far right. Sato et al. (2012) suggest that the reason for this is that on-cells are strongly activated by both the background textures in the upper and lower leftmost stimuli; as the texture elements are made denser, the differential activation of on-cells versus off-cells is weakened. A density effect of this sort may have undermined the efforts of Solomon, Sperling, and Chubb (1993), who addressed this same issue but found that textures designed to

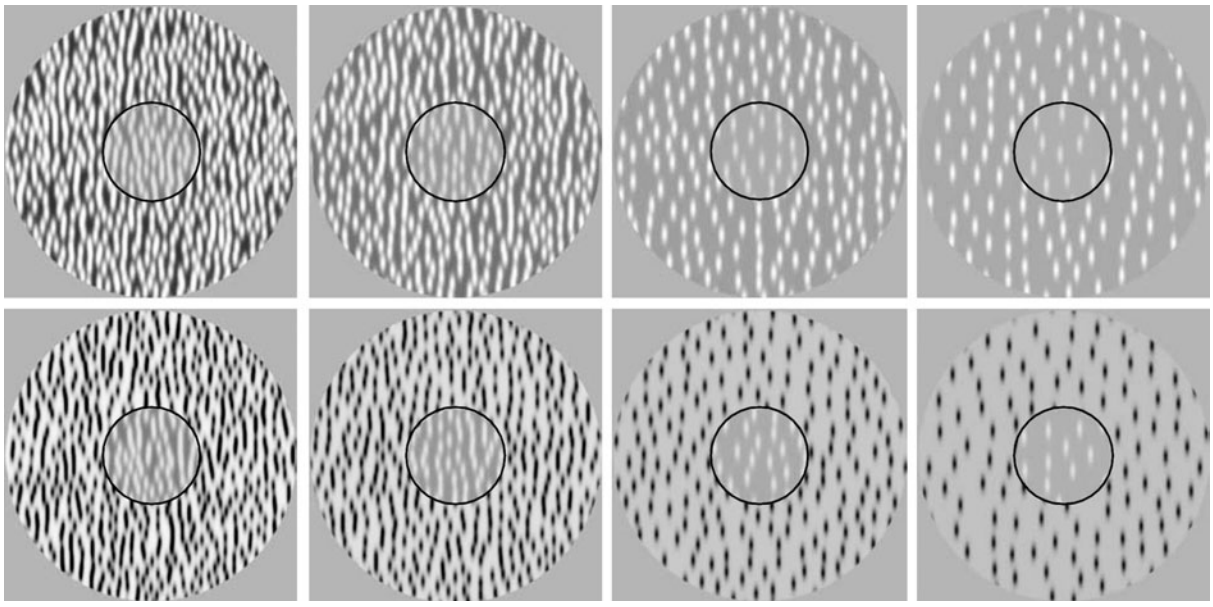


Figure III.41-5. Is the contrast contrast illusion selective for contrast polarity? Most observers judge the test disks in the two rightmost lower panels to be higher in contrast than the corresponding test disks in two rightmost upper panels. This suggests that the contrast contrast illusion is stronger if the polarity of the background texture matches that of the target texture. (Figure reproduced with permission from *Journal of Vision*.)

differentially activate on-cells versus off-cells were equally effective at suppressing the contrast of each other.

COLOR AND THE CONTRAST CONTRAST ILLUSION

It is possible to create powerful contrast contrast effects using displays in which texture varies across space not in light intensity but rather in color. An example is shown in Figure III.41-1b. The textures used in this figure are approximately uniform in their luminance (their effective light intensity for human vision); instead, they vary in their chromatic contrast. The background texture on the right varies strongly between a red and a green that are roughly equiluminant, and the texture patch reproduced on the left and right sides of the display varies less dramatically. Most observers perceive the texture disk on the left to be higher in chromatic contrast than the disk on the right.

Contrast contrast effects have been studied extensively by Singer and D'Zmura (1994) using textures modulating not only achromatic light intensity but also color. They used textures that modulate between high and low luminance as well as textures that modulate between equal intensity (equiluminant) red and green and other textures that modulate between equal intensity blue and yellow. They tested how strongly each of these types of texture suppresses the perceived contrast (either color contrast or intensity contrast) of a central test patch of each type. Their studies revealed that each type of texture is most effective at suppressing the contrast of a test patch of the same type of texture; however, substantial suppression also occurs across different types of texture (with the exception that equiluminant color modulations exert very little influence on the perceived contrast of achromatic texture).

ECOLOGICAL ACCOUNTS OF THE CONTRAST CONTRAST ILLUSION

Readers may be struck by the sense that the right side of Figure III.41-1a appears as though they are looking at high-contrast texture through a translucent plastic disk. Lotto and Purves (2001) and Purves, Williams, Nundy, and Lotto (2004) have argued that the contrast contrast illusion is much stronger when the visual features of the stimulus are consistent with such an interpretation than when they are not. Consider Figure III.41-1c, for example. The left and right sides in this figure are identical except that the background on the left side has been contrast-reversed: wherever the right-side background is white, the left-side background is black. The right side is designed so that the texture appears to run continuously behind a translucent plastic disk in the middle. This effect is undermined, however, on the left side: if an image region abutting one side of the disk boundary is bright, then it switches to dark on the other side. And it is undoubtedly true that to most observers the central patch on the left of Figure III.41-1c appears higher in contrast than the identical patch on the right.

Lotto and Purves (2001) and Purves et al. (2004) observe that if indeed the image projected to the retina were the result of seeing high-contrast texture through a translucent disk, then the ways in which light interacts with surfaces would mandate that the contrast of the image region subtended by the disk should be low. However, if instead the image were the result of light reflecting directly from a textured surface whose reflectance properties changed at the disk boundary, then no constraints would operate to restrict the contrast of the texture inside the disk. They argue that, in the automatic processing it performs prior to any conscious reflection on the part of the observer, vision comes to embody this knowledge about the probable causes of the image projected to the retina. Why does the disk on the left side appear lower in contrast? It is because the features in the image make it likely that the distal cause of the image is a translucent disk through which high-contrast texture is being seen (Lotto & Purves, 2001; Purves & Lotto, 2003; Purves et al., 2004).

If Lotto and Purves (2001), Purves and Lotto (2003), and Purves et al. (2004) are correct, then models based strictly on lateral inhibition are unlikely to suffice to account for the contrast contrast illusion because such models do not take into account the full scope of stimulus features that reliably signal that the stimulus results from viewing a textured background through a translucent disk. Some of the displays they have created offer striking support for this claim. It is important to realize, however, that regardless of whether this theory is true, one is still left with the problem of modeling the computations that occur within the visual system that give rise to the contrast contrast illusion.

REFERENCES

Barch, D. M., Carter, C. S., Dakin, S. C., Gold, J., Luck, S. J., MacDonald, A. III, . . . Strauss, M. E. (2012). The clinical translation of a measure of gain control: the contrast-contrast effect task. *Schizophrenia Bulletin*, *38*, 135–143.

Bosten, J. M., & Mollon, J. D. (2010). Is there a general trait of susceptibility to simultaneous contrast? *Vision Research*, *50*, 1656–1664.

Cannon, M. W., & Fullenkamp, S. C. (1993). Spatial interactions in apparent contrast: Individual differences in enhancement and suppression effects. *Vision Research, 33*(12), 1685–1695.

Chubb, C., Sperling, G., & Solomon, J. (1989). Texture interactions determine perceived contrast. *Proceedings of the National Academy of the Sciences USA*, *86*, 9631–9635.

Dakin, S., Carlin, P., & Helmsley, D. (2005). Weak suppression of visual context in chronic schizophrenia. *Current Biology*, *15*(20), R822–R824.

Gold, J., Barch, D. M., Carter, C. S., Dakin, S. C., Luck, S. J., MacDonald, A. III, . . . Strauss, M. (2012). Clinical, functional, and intertask correlations of measures developed by the Cognitive Neuroscience Test Reliability and Clinical Applications for Schizophrenia Consortium. *Schizophrenia Bulletin*, *38*, 144–152.

Heeger, D. J. (1994). The representation of visual stimuli in primary visual cortex. *Current Directions in Psychological Science, 3*(5), 159–163.

Hubel, D. H. (1960). Single unit activity in lateral geniculate body and optic tract of unrestrained cats. *Journal of Physiology, 150*, 91–104.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, *160*, 106–154.

Kuffler, S. W. (1953). Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology, 16*, 37–68.

Lotto, R. B., & Purves, D. (2001). An empirical explanation of the Chubb illusion. *Journal of Cognitive Neuroscience*, *13*(5), 547–555.

Mach, E. (1866). Über den physiologischen Effect raümlich vertheilterLichtreize: II. Sber. math.-nat [On the physiological effect of spatially distributed luminance stimuli]. *Classe kaiserlichen Akademie der Wissenschaften, 54*, 131–144.

Purves, D., & Lotto, R. B. (2003). *Why we see what we do: An empirical theory of vision*. Sunderland, MA: Sinauer Associates.

Purves, D., Williams, S. M., Nundy, S., & Lotto, R. B. (2004). Perceiving the intensity of light. *Psychological Review, 111*(1), 142–158.

Sato, H., Motoyoshi, I., & Sato, T. (2012). Polarity selectivity of spatial interactions in perceived contrast. *Journal of Vision 12*(2), 1–10. http://www.journalofvision.org/content/12/2/3

Singer, B., & D'Zmura, M. (1994). Color contrast induction. *Vision Research, 34*(23), 3111–3126.

Singer, B., & D'Zmura, M. (1995). Contrast gain control: A bilinear model for chromatic selectivity. *Journal of the Optical Society of America A, 12*(4), 667–684.

Solomon, J. A., Sperling, G., & Chubb, C. (1993). The lateral inhibition of perceived contrast is indifferent to on-center/off-center segregation, but specific to orientation. *Vision Research, 33*(18), 2671–2683.

Yang, E., Tadin, D., Glasser, D. M., Wook Hong, S., Blank, R., & Park, S. (2012). Visual context processing in schizophrenia. *Clinical Psychological Science*. Advance online publication. http://cpx.sagepub.com/content/1/1/5.full

Chapter 42
Illusory Color Spread from Apparent Motion

Carol M. Cicerone and Donald D. Hoffman

INTRODUCTION

The visual system is capable of constructing illusory colors and contours that may be absent in the physical stimulus (e.g., Grosof, Shapley, & Hawken, 1993; Kanisza, 1979; Michotte, Thines, & Crabbe, 1964; Nakayama & Shimojo, 1990, 1992; Nakayama, Shimojo, & Ramachandran, 1990; Peterhans & von der Heydt, 1991; van Tuijl, 1975; Varin, 1971; Yamada, Fujita, & Masuda, 1993). In particular, motion is effective in allowing the visual system to use multiple fragmented views of an object over time to reconstruct its shape as a whole (e.g., Andersen & Braunstein, 1983; Andersen & Cortese, 1989; Gibson, 1979; Kaplan, 1969; Lappin, Doner, & Kottas, 1980; Shipley & Kellman, 1983, 1984; Stappers, 1989; Wallach & O'Connell, 1953; Wertheimer, 1923; Yonas, Craton, & Thompson, 1987). We introduced an effect called *color from motion* in which the perception of apparent motion is accompanied by the perception of illusory color that is seen in physically achromatic regions of the stimulus (Cicerone & Hoffman, 1992). We explored this phenomenon to define the conditions under which it occurs (Cicerone, Hoffman, Gowdy, & Kim, 1995; Miyahara & Cicerone, 1997); to understand how it might be useful to organize the visual scene (Chen & Cicerone, 2002a); and to link it to the perception of motion, contour, color, and depth (Chen and Cicerone, 2002b).

A typical display of color from motion is shown in Figure III.42-1. Each frame (8° of visual angle on a side) consisted of an achromatic background field (CIE $x = 0.276$, $y = 0.286$, luminance 73 cd/m^2) over which was randomly arrayed 800 to 1,200 dots (each 3.5 min of arc in diameter). Within a circular test region (typically 1° to 2° of visual angle in diameter), the dots were colored green (CIE $x = 0.280$, $y = 0.610$). All other dots were red (CIE $x = 0.621$, $y = 0.344$). To create successive frames, no dots changed their locations within the frame; only the color assignments of the dots were changed, according to a uniform vertical displacement (0.12° of visual angle) of the test region in successive frames. When frames are cycled, typically with an effective displacement rate of the test region equivalent to 7°/s, up and down over a vertical region spanning 5° of visual angle, an illusory green disk moving up and down pops into view and a spread of illusory green color is seen in the physically achromatic test region. This effect and others described in this chapter can be viewed in the online publication of Chen and Cicerone (2002a).

The rating methods used in the early studies (Cicerone & Hoffman, 1992; Cicerone et al., 1995) established the salience of the color spreading effect and its link to the perception of apparent motion. Certain aspects of the illusory color, however, such as its saturation, require more sensitive methodologies. Miyahara and Cicerone (1997) used a side-by-side matching method and found that the hue of the subjective color spread approximates that of the test dots. In their methodology, the matching stimulus was stationary and homogeneously colored, whereas the color from motion stimulus was perceived as moving and included test and surround dots. Chen and Cicerone (2002a, b) used real lights to cancel the subjective color spread in color from motion, while keeping luminance constant, to measure the hue of the illusory color spread. An increase in the luminance of the test dots produced an increase in the saturation of the physical lights required to match the subjective color spread as measured by the side-by-side method (Miyahara & Cicerone, 1997) and the cancellation method (Chen & Cicerone, 2002a, b).

PERCEPTION OF MOTION IS ESSENTIAL IN COLOR FROM MOTION

Illusory color spread as measured by a rating method is linked to apparent motion of the test region. Therefore we asked if the salience of the color spread as measured with cancellation was linked to the salience of the apparent motion (Chen & Cicerone, 2002a, b). We varied the effective rate of vertical translation of the test region between zero and 12° of visual angle per second. The results for two observers (Fig. III.42-2) can be well described by two linear functions with a steeply sloping first branch (from zero to 1°/s) and a second branch with a much reduced, near-zero slope (from 1°/s to 12°/s). The general profile of the results suggests an all-or-none relationship between apparent motion and illusory color spread. Observers report that as the translation speed of the test region increases, the perception of the illusory, green-colored disk as a separate form that moves over the field of dots is enhanced. This perception of separation is reported reliably for speeds exceeding 1°/s. Concurrently, the test dots appear to assume the color of the surround dots so that all dots appear to have the same color, in this case red, even if the test dots are

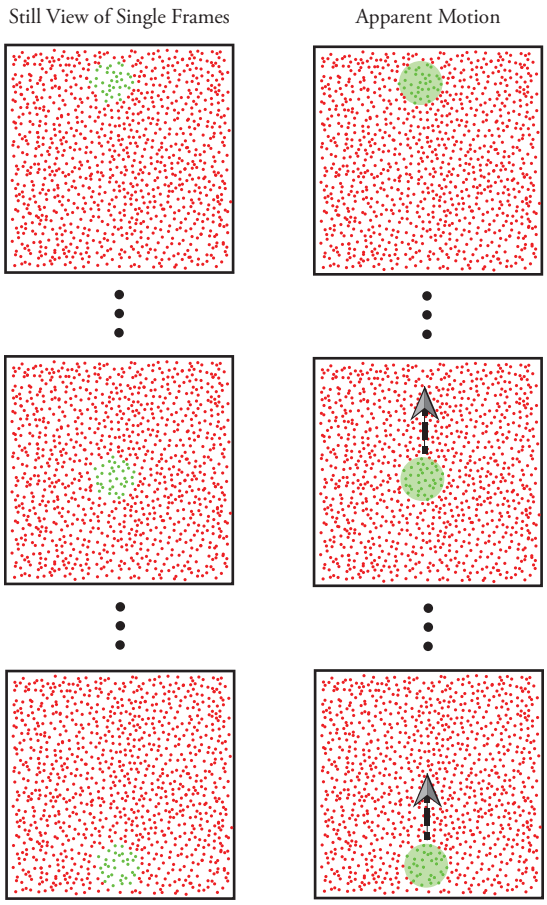


Figure III.42-1. A typical display of color from motion. Each frame (8° of visual angle on a side) consists of an achromatic background field (CIE $x = 0.276$, $y = 0.286$, luminance 73 cd/m^2) over which was randomly arrayed 800 to 1,200 dots (each 3.5 min of arc in diameter). Within a circular test region (typically 1° to 2° of visual angle in diameter), the dots were colored green (CIE $x = 0.280$, $y = 0.610$). All other dots were red (CIE $x = 0.621$, $y = 0.344$). The luminance of the red and green dots could be independently varied. To create successive frames, no dots changed their locations within the frame; only the color assignments of the dots were changed, according to a uniform vertical displacement (0.12° of visual angle) of the test region in successive frames. Left: Still views of successive frames. Right: When frames are cycled, typically with an effective displacement rate of the test region equivalent to 7°/s, up and down over a vertical region spanning 5° of visual angle, an illusory green disk moving up and down pops into view and a spread of illusory green color is seen in the physically achromatic test region. (Image: Carol M. Cicerone & Donald D. Hoffman.)

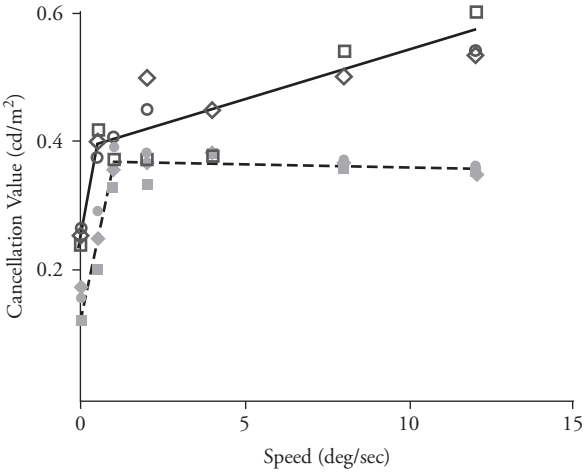


Figure III.42-2. The salience of the color spread as measured with cancellation is linked to the salience of the apparent motion. We varied the effective rate of vertical translation of the test region between zero and 12° of visual angle per second. The results for two observers can be well described by two linear functions with a steeply sloping first branch (from an effective speed of translation from zero to 1°/s) and a second branch of much reduced or zero slope (from 1°/s to 12°/s). (Adapted from Chen & Cicerone, 2002a.)

physically colored green. Thus it appears that for speeds of effective translation of the test region that are greater than 1°/s there is little or no enhancement of the illusory color spread whereas the salience of the perceived separation of the figure—defined by the illusory color spread—from the array of dots is enhanced.

COLOR FROM MOTION WITHOUT CONTOUR FORMATION

Is color from motion linked to contour formation as well as to the perception of motion? Luminance contrast is necessary for the formation of static subjective contours (e.g., Frisby & Clayworthy), for the perception of apparent motion in achromatic stimuli (e.g., Cavanagh, Boeglin, & Favreau, 1985; Ramachandran & Gregory, 1978), and for the perception of achromatic neon spreading (e.g., Bressan, 1993). If, near equiluminance, color spread occurs without the formation of a subjective contour, then color from motion is likely to be regulated by mechanisms distinct from those regulating contour formation. Miyahara and Cicerone (1997) tested this idea with green (test) and red (surround) dots that were matched in luminance as determined for each observer individually by means of heterochromatic flicker photometry. Chromaticity differences between the test dots and the surround dots in the absence of luminance differences produced the perception of spread of color from motion. As the luminances of the test and surround dots approached equality, the strength of the subjective contour surrounding the test region was reduced while the subjective color spread was still perceived. Observers reported that in such equiluminant conditions, there was no clear contour bounding the region of the subjective color spread and that apparent motion was "not smooth" and "slower" than the conditions in which the test and surround dots differed in luminance (Chen & Cicerone, 2002a). These results are consistent with findings that suggest that the neural mechanisms responsible for contour formation rely on luminance information (e.g., Kanizsa, 1979; Marr, 1982). These findings support the idea that the mechanisms regulating color from motion are separable from those regulating contour formation and that color spread can occur without contour formation.

Another way to study the role of luminance differences is to use red–green dichromatic observers who are incapable of making red–green discriminations on the basis of chromaticity discriminations alone. Miyahara and Cicerone (1997) presented these stimuli to a deuteranope. The deuteranope

lacks the middle wavelength sensitive pigment and is therefore insensitive to chromaticity differences in the middle- to long-wavelength range of the visible spectrum. In this range, where our red and green stimuli lie, the deuteranope sees only luminosity differences, based on the activity of the long wavelength sensitive cones. Thus the results of a red–green dichromat should allow us to assess the impact of luminosity differences alone for these stimuli. When the luminosity of the green dots in the test region was high relative to that of the dots in the surround, the deuteranope saw a bright disk moving over the test region. As expected, near equiluminance between the test and surround dots, the deuteranope in our study saw no apparent motion, nor did he see brightness spread in the test region. This is in contrast to color-normal observers who saw apparent motion and color spread without a clearly defined contour when test and surround dots were matched in luminance.

MODAL VERSUS AMODAL COMPLETION IN COLOR FROM MOTION

There are two modes in which color from motion is perceived: (1) as a localized change of illumination, a colored spotlight or shadow, moving over a textured surface (Cicerone & Hoffman, 1992; Cicerone et al., 1995; Miyahara & Cicerone, 1997) or (2) as a moving, colored object seen through apertures in an occluding surface (Chen & Cicerone, 2002a, b; Cicerone & Hoffman, 1997). The mode in which color from motion is seen depends on figural cues and regional differences in the luminance contrast between the chromatic elements and the achromatic background. Regions with figural cues tend to be seen as moving. These regions are seen in the first mode (modal completion; Fig. III.42-3, left) if their defining figural elements, dots in this case, are of lower luminance as compared to the background and in the second mode (amodal completion, Fig. III.42-3, right) if the defining figural elements are of higher luminance as compared to the background. As compared with color from motion seen in modal completion, the perceived color in amodal completion is markedly higher in saturation and the organization of the scene is different in that objects are perceived to lie behind a partially occluding screen. Nonetheless, in both cases, the perception of motion is linked to a spread of color over regions defined by motion. Hence the mechanisms driving the color spread, whether a desaturated veiling color (modal completion) appearing to glide over the display or a highly saturated disk (amodal completion) moving behind a partially occluding screen, are likely to be the same.

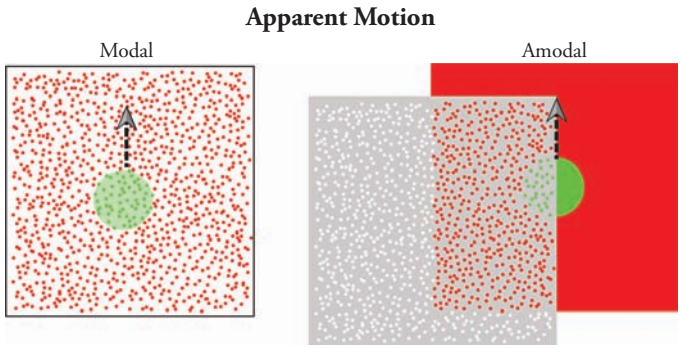


Figure III.42-3. There are two modes in which color from motion can be perceived, either in modal completion (left) as a localized change of illumination, a colored spotlight, or a shadow, moving over a textured surface, or in amodal completion (right) as a moving, colored object seen through apertures in an occluding surface. The mode in which color from motion is seen depends on figural cues and regional differences in the luminance contrast between the chromatic elements and the achromatic background. Regions with figural cues tend to be seen as moving. These regions are seen in modal completion if their defining figural elements (dots in this case) are of lower luminance as compared to the background and in amodal completion if the defining figural elements are of higher luminance as compared to the background. (Image: Carol M. Cicerone & Donald D. Hoffman.)

LUMINANCE RELATIONSHIPS SHAPE VISUAL SCENE ORGANIZATION IN COLOR FROM MOTION

In still view of our standard stimulus, the small test region of green dots is clearly seen as the figure, and the surrounding region of red dots is seen as the ground. Regardless of the luminance of the achromatic background, the test region is seen to move, and color spread is linked to this moving region. To explore the importance of figural cues, Chen and Cicerone (2002a) used stimuli composed of alternating bands of equal widths of red and green dots whose luminance contrasts, as compared to the achromatic background, could be manipulated. For such stimuli without clear figure–ground configurations, we asked whether color from motion remains a salient effect and how luminance relationships help to organize the visual scene. Regardless of the luminance of the achromatic background, apparent motion and color spread are associated with regions of lower luminance contrast. A switch in perception of seeing red bands moving or green bands moving occurs at each observer's point of equiluminance between the red and green stimuli. The relative widths of the bands of red and green dots were varied to test whether figure–ground cues could supersede luminance cues. Indeed, for bands that were thin enough (roughly an 8:1 ratio for observers), the narrower bands (more figure-like) were seen as moving regardless of luminance contrast relationships (Chen & Cicerone, 2002b).

COLOR FROM MOTION IS REGULATED AT A POINT BEYOND BINOCULAR COMBINATION

The dependence of color spread on the perception of apparent motion in this phenomenon and the spread of color in the absence of contour formation suggest that the locus of the mechanisms underlying color from motion may be at a point beyond binocular combination. Evidence supporting this idea was obtained by the dichoptic presentation of every other frame of the full stimulus sequence to one eye and, out of phase, the alternate frames to the other eye. To overcome binocular rivalry, the eye that did not receive the standard stimulus was presented with a stimulus that was identical in every way, except that the test dots were absent. A compelling perception of color from motion,

as measured by a rating method, was seen that was equal to that obtained when the full stimulus sequence was presented to each eye alone (Cicerone & Hoffman, 1997). This is consistent with the regulation of color from motion at sites beyond the convergence of monocular pathways.

Perhaps even more persuasively, Chen and Cicerone (2002b) showed that depth, as well as form and color, is recovered from apparent motion. Stimuli were presented dichoptically with left- and right-eye views identical as to the *locations* of all dots. Binocular disparity was introduced by means of translations in the *color assignments* for corresponding image elements in the two eyes. Horizontal crossed or uncrossed disparities of 0.5° of visual angle were created by differences in color assignments alone. In random presentation of the crossed or uncrossed horizontal displacements, observers were asked to judge if the illusory figure defined by the color spread was *behind* or *in front of* the field of dots. In still view of the stimulus, binocular rivalry occurred and neither apparent motion, nor color spread, nor depth was seen. In this case, observers performed at chance level when required to judge depth. When the frames were cycled at an effective rate of 7°/s over a vertical distance upward and downward spanning 10° of visual angle, apparent motion, color spread, and depth were perceived. In random presentations of stimuli for crossed and uncrossed disparities in motion view, observers performed well above chance (95% confidence interval) with some observers performing with 100% accuracy.

GENERAL DISCUSSION

The Distinctive Features of Color From Motion

The effect we call color from motion is distinctive in a number of ways. First, neither contour formation nor color spread is seen in still view of our stimuli. In this way it is distinct from static neon color spreading, an effect that is well known. Furthermore, illusory color spread as seen in color from motion is not a general feature in motion stimuli; for example, it is not reported in kinetic occlusion. Second, in color from motion there are no spatial displacements of the dots; only the color assignments of the dots change from frame to frame. Apparent motion and the attendant illusory color spread are generated only by the change in chromaticity or luminance of the dots. To buttress this second point, we created stimuli in which the test region remained fixed in space and the test dots were set in motion either (a) independently and randomly, (b) in unison along the same trajectory, or (c) in unison along a random trajectory. Naïve observers were tested with all of these stimuli, and none reported seeing color spread (Chen & Cicerone, 2002a). Third, although the saturation of the illusory color spread increases with increases in the luminance of the *test* dots, the luminance of the dots in the *surround* region has no impact on the illusory color spread when the chromaticities of the test and surround dots differ (Chen & Cicerone, 2002a). This differentiates color from motion from color contrast, wherein the luminance of surround elements has considerable impact. Fourth, subjective color spread was seen without the perception of a subjective contour near the point of equiluminance between test and surround dots, as long as there was a chromaticity difference between the dots (Miyahara & Cicerone, 1997). This result suggests that the spread of illusory color in color from motion does not require the prior formation of a contour and that color, independent of contour, can be recovered in tandem with seeing motion.

When color from motion is seen in modal completion, the low saturation and neon-like quality of the illusory color spread is reminiscent of the quality of the perception in displays of transparency. Is color from motion the same as transparency? We argue that it is not for the following reasons. First, the perception of transparency occurs in displays due to both figural and luminance cues that are already present in the stimulus (e.g., Adelson, 1993; Beck, 1978; da Pos, 1989; D'Zmura, Colantoni, Knoblauch, & Laget, 1997; Metelli, 1974), whereas in color from motion a new colored surface, with or without a border, is created by the visual system in physically achromatic regions. In other words, when transparency is perceived, physically present but differentiated regions are unified into a single perceptual layer, whereas in color from motion, an entirely new, illusory layer is constructed by the visual system. Second, motion is not required for transparency to be perceived, whereas color from motion requires the perception of apparent motion and is never seen in still view. Third, color from motion can be seen in amodal completion (Chen & Cicerone, 2002a; Cicerone & Hoffman, 1997), a perception that differs markedly from any of the characteristics of transparency.

Our results indicate that color from motion is regulated at a point beyond binocular combination (Cicerone & Hoffman, 1997); that it requires the concurrent, if not prior, perception of motion (Chen & Cicerone, 2002a; Cicerone et al., 1997); that, in addition to form and color, depth can be recovered in color from motion (Chen & Cicerone 2002b); and that figure–ground configuration can override luminance relationships as the determinant of which areas appear to move and to be filled with illusory color (Chen & Cicerone, 2002a). Considering our current understanding of visual processing in the primate brain, these findings suggest that the mechanisms supporting the perception of color from motion include neural processing at higher levels.

The Functional Significance of Color From Motion

Can our results be related to the visual system's ability to break visual camouflage? In natural scenes, objects or surfaces may not be perceived because of the color, luminance, or texture of nearby surfaces. We reasoned that if color from motion is a robust camouflage-breaking mechanism, then it should be able to render the test object visible even when color is not an obvious cue. The stimulus was modified so that a proportion (zero to 0.8) of the dots in the surround region were green instead of all red. In still view, the test region with green dots was not reliably seen; thus, in still view, the test region was effectively camouflaged. Nonetheless, when the stimulus sequence was cycled as before and apparent motion was perceived, a moving, illusory green disk was seen (Cicerone & Hoffman, 1997).

In other natural scenes, objects may be hidden from full view by occlusion. To mimic this situation, we reduced the illumination of the achromatic background in our stimuli to show that color from motion can be seen in amodal completion as a highly saturated green disk that moves over a highly saturated red background, all seen through random perforations in a dark screen (Chen & Cicerone, 2002a; Cicerone & Hoffman, 1997). The mode in which color from motion is seen depends on figural cues and on regional differences in luminance contrast between the chromatic elements and the achromatic background (Chen & Cicerone 2002a).

In still view, the physical representation of the scene may give an equivocal interpretation of objects and surfaces. When the test is seen in apparent motion, subjective color spread helps to reveal the hidden object in modal or in amodal completion. Furthermore, not only form and color but also depth can be recovered in tandem with seeing motion. We propose that the neural mechanisms that support perceptions in *color from motion* may be the same as those that work in natural scenes to reveal form, color, and depth to the visual system, even when it is confronted with fragmented physical information, as occurs in camouflage.

REFERENCES

Adelson, E. H. (1993). Perceptual organization and the judgment of brightness. *Science*, *262*, 2042–2044.

Andersen, G. J., & Braunstein, M. L. (1983). Dynamic occlusion in the perception of rotation in depth. *Perception & Psychophysics*, *34*, 356–362.

Andersen, G. J., & Cortese, J. M. (1989). 2-D contour perception resulting from kinetic occlusion. *Perception & Psychophysics*, *46*, 49–55.

Beck, J. (1978). Additive and subtractive color mixture in color transparency. *Perception & Psychophysics*, *23*, 265–267.

Bressan, P. (1993). Revisitation of the luminance conditions for the occurrence of the achromatic neon color spreading illusion. *Perception & Psychophysics*, *54*, 55–64.

Cavanagh, P., Boeglin, J., & Favreau, O. E. (1985). Perception of motion in equiluminous kinematograms. *Perception*, *14*, 151–162.

Chen, V. J., & Cicerone, C. M. (2002a). Subjective color from apparent motion. *Journal of Vision*, *2*, 424–437. http://jov.arvojournals.org/article.aspx?articleid=2121458

Chen, V. J. & Cicerone, C. M. (2002b). Depth from subjective color and apparent motion. *Vision Research*, *42*, 2131–2135.

Cicerone, C. M., & Hoffman, D. D. (1992). Dynamic neon colors: Perceptual evidence for parallel visual pathways. *Advances in Color Vision*, *Technical Digest, 4*, 66–68.

Cicerone, C. M., & Hoffman, D. D. (1997). Color from motion: Dichoptic activation and a possible role in breaking camouflage. *Perception*, *26*, 1367–1380.

Cicerone, C. M., Hoffman, D. D., Gowdy, P. D., & Kim, J. S. (1995). The perception of color from motion. *Perception & Psychophysics*, *57*, 761–777.

da Pos, O. (1989). *Trasparenze*. Padua, Italy: Icone.

D'Zmura, M., Colantoni, P., Knoblauch, K., & Laget, P. (1997). Color transparency. *Perception*, *26*, 471–492.

Frisby, J. P. & Clayworthy, J. L. (1975). Illusory contour: Curious cases of simultaneous brightness contrast? *Perception*, *4*, 349–357.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Grosof, D. H., Shapley, R. M., & Hawken, M. J. (1993). Macaque V1 neurons can signal "illusory" contours. *Nature*, *365*, 550–552.

Kanizsa, G. (1979). *Organization in vision*. New York: Praeger.

Kaplan, G. A. (1969). Kinetic disruption of optical texture: The perception of depth at an edge. *Perception & Psychophysics, 6*, 193–198.

Lappin, J. S., Doner, J. F., & Kottas, B. (1980). Minimal conditions for the visual detection of structure and motion in three dimensions. *Science*, *209*, 717–719.

Marr, D. (1982). *Vision*. New York: Freeman.

Metelli, F. (1974). The perception of transparency. *Scientific American*, *230*, 90–98.

Michotte, A., Thines, G., & Crabbe, G. (1964). *Amodal completion of perceptual structures*. Louvain, France: Publications Universitaires de Louvain.

Miyahara, E., & Cicerone, C. M. (1997). Color from motion: Separate contributions of chromaticity and luminance. *Perception*, *26*, 1381–1396.

Nakayama, K., & Shimojo, S. (1990). Toward a neural understanding of visual surface representation. *Cold Spring Harbor Symposia on Quantitative Biology*, *55*, 911–924.

Nakayama, K., & Shimojo, S. (1992). Experiencing and perceiving visual surfaces. *Science*, *257*, 1357–1363.

Nakayama, K., Shimojo, S., & Ramachandran, V. S. (1990). Transparency: Relation to depth, subjective contours, luminance and neon color spreading. *Perception*, *19*, 497–513.

Peterhans, E., & von der Heydt, R. (1991). Subjective contours: Bridging the gap between psychophysics and physiology. *Trends in Neurosciences*, *14*, 112–119.

Shipley, T. F., & Kellman, P. J. (1994). Spatiotemporal boundary formation: Boundary, form, and motion perception from transformations of surface elements. *Journal of Experimental Psychology: General*, *123*, 3–20.

Stappers, P. J. (1989). Forms can be recognized from dynamic occlusion alone. *Perceptual and Motor Skills*, *68*, 243–251.

van Tuijl, H. F. (1975). A new visual illusion: Neon-like color spreading and complementary color induction between subjective contours. *Acta Psychologica*, *39*, 441–445.

Varin, D. (1971). On phenomena relating to contrast and chromatic diffusion in the spatial organization of perceptual space. *Rivista di Psicologia*, *65*, 101–128.

von der Heydt, R., & Peterhans, E. (1989). Mechanisms of contour perception in monkey visual cortex: I. Lines of pattern discontinuity. *Journal of Neuroscience*, *9*, 1731–1748.

Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, *45*, 427–430.

Yamada, W., Fujita, N., & Masuda, N. (1993). Amodal completion as another perception of color spreading stimuli. *Perception and Motor Skills*, *76*, 1027–1033.

Yonas, A., Craton, L. G., & Thompson, W. B. (1987). Relative motion: Kinetic information for the order of depth at an edge. *Perception & Psychophysics*, *41*, 53–59.

Chapter 43
The Reversed Contrast Necker Cube

Alessandra Galmonte and Tiziano Agostini

PRESENTATION AND DESCRIPTION OF THE EFFECT

It is well known that a grey surface surrounded by a light background appears darker than an identical grey surface surrounded by a dark background. This phenomenon is called simultaneous contrast (see Fig. III.43-1).

In our reversed contrast Necker cube display (Agostini & Galmonte, 2002), two middle grey dashed cubes were used. The first one, which had dark inducer corners, was placed on a light inducer background, and the second one, which had light inducer corners, was placed on a dark inducer background.

We also tested two control conditions, identical to the two experimental conditions except for (a) the colors of the inducer corners, which were the same middle grey as the target regions or (b) disks with the same area and luminance replaced the inducer corners. Observers reported that the grey belonging to the dark inducer corners appeared lighter and that the grey belonging to the light inducer corners appeared darker (see Fig. III.43-2). As predicted, in both the control experiments neither of these situations occurred (see Figures III.43-3 and III.43-4).

WHY THE EFFECT IS INTERESTING AND RELEVANT FOR VISION RESEARCH

We used this novel display of simultaneous contrast with the aim, for the first time, to directly compare local contrast and perceptual belongingness. On the one hand, lateral inhibition (or local contrast) models (Cornsweet, 1970; Hurvich & Jameson, 1957) claim that, in the simultaneous contrast effect, the receptors stimulated by the light background send inhibition to the receptors stimulated by the grey surface, causing its perceptual darkening. However, the receptors stimulated by the dark background send little inhibition, and, therefore, there is not a darkening effect.

On the other hand, simultaneous contrast phenomena have been explained according to principles of perceptual organization (Gilchrist et al., 1999; Wertheimer, 1923); that is, the perceived color of a surface is mainly determined by the global contrast with the other colored surfaces to which it perceptually belongs and not simply by low-level processes of local contrast involving lateral inhibition.

The term "perceptual belongingness" refers to the result of a connection process that unifies the specific perceptual attributes arising from spatial articulation, whereas "perceptual spatial articulation" refers to the result of a perceptual organization process of selecting and highlighting specific physical spatial relationships of a visual pattern.

Consider the following example. A number of identical small disks are equidistantly spaced along the perimeter of a circle. Perceptual organization processes acting on the physical attributes of this pattern will give rise to perception of a circle of small disks. This is a result of perceptual spatial articulation. Each individual disk appears to belong to the circle. It is in this way that perceptual spatial articulation determines the specific belongingness relationships among the elements (Agostini & Galmonte, 1999).

We decided to use two middle-grey dashed cubes, one having dark inducer corners placed on a light inducer background and the other having light inducer corners placed on a dark inducer background, because, according to the local contrast explanations, the effect is determined only by the influence of the immediate surrounding regions, therefore, the middle grey must be contrasted only by the inducer backgrounds, favored also by the dashing. Observers were asked to perceive the cube as a whole and then to perceptually reverse it a couple of times before judging the lightness of the entire middle-grey dashed lines. This process guaranteed a global judgment rather than a local one. Then they were asked to select from a Munsell scale for the judgment of lightness, placed on the same background of the tested displays, the shade of grey that was more similar to that of the dashed lines.

Placing the scale on the same background guaranteed that any local-contrast effects of the background on the scale elements would be the same as any local-contrast effects of the background on the elements of the test display. This new method can therefore reveal high-level effects that act beyond the effects of the surround. This is similar to setting the value of the measurement instrument to zero (target regions and scale are both affected by the same surround). Therefore, zero represents the magnitude of the local contrast effect, which obviously should correspond to the objective luminance value of the target region. With this method it became possible to obtain a relative measure of the strength of belongingness. The result is that judgments reflected the strength of global factors only.



Figure III.43-1. The classical simultaneous lightness contrast phenomenon. The grey target surrounded by white appears darker than a physically identical grey target surrounded by black (Image by Public Domain).

Therefore, according to the belongingness explanation, the middle grey must be contrasted also by the color of the inducer corners. In fact, the dashed lines are necessary to complete the cube, therefore they belong to the corners, which, being of a different color, should affect the perceived color of the dashed lines.

Furthermore, if local contrast prevails then the middle-grey dashed lines of the cube on the light background should be perceived as darker than the middle-grey dashed lines of the cube placed on the dark background. Meanwhile, if belongingness prevails, the perceptual outcome should be in the opposite direction.

In the first control experiment, in which the dashed lines and the corners were the same color, we predicted there would be no effect. In the second control experiment, we controlled also for the luminance of the inducer corners. Under these conditions, the percept was that of a dashed cube that had all of its corners occluded by disks. In this case, the colored disks should not affect the color of the dashed lines because the lines do not belong to the disks but amodally complete a cube behind them.

Our results are very relevant, since we demonstrated for the first time that when higher-level factors act contemporaneously with lower-level factors, the contrast effect induced by the global organization principle of perceptual belongingness prevails on retinal lateral inhibition.

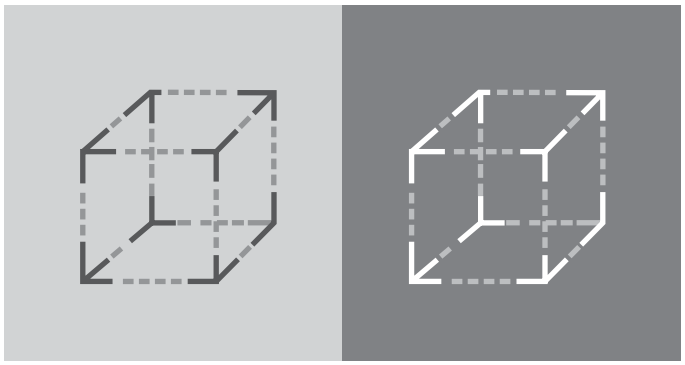


Figure III.43-2. Even though the grey dashed lines of the Necker cube represented on the left side of the figure are completely surrounded by white, they appear lighter since they are perceived to belong to the dark inducer corners, and vice versa for the right figure (Image by Revised from [Agostini, T., & Galmonte, A. (2002). Perceptual organization overcomes the effect of local surround in determining simultaneous lightness contrast. *Psychological Science, 13*(1), 88–92]).

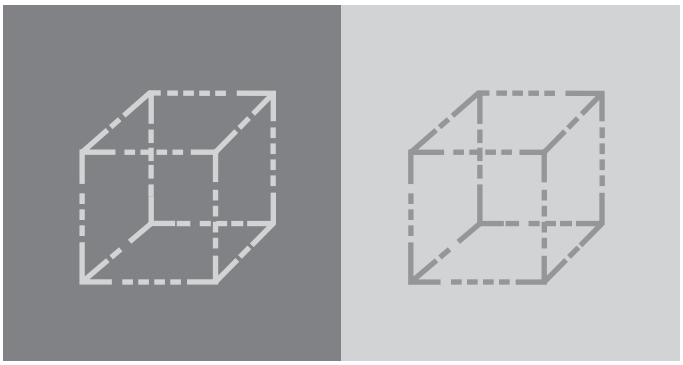


Figure III.43-3. When the dashed lines and the corners are made of the same shade of grey, the induction effect goes in the direction of local contrast (Image by Revised from [Agostini, T., & Galmonte, A. (2002). Perceptual organization overcomes the effect of local surround in determining simultaneous lightness contrast. *Psychological Science, 13*(1), 88–92]).

VARIANTS AND RELEVANT PARAMETERS FOR THE EFFECT

We also successfully tested whether global grouping factors induction can be found in the chromatic domain as well (Agostini, Da Pos, Galmonte, & Mauri, 2009). Our Necker cube display was reproduced using red/yellow-inducing colors (backgrounds/corners) and orange-induced dashed lines. There were also two control conditions, as in the achromatic version. The procedure was the same.

Results replicate those obtained for the achromatic version. In fact, global factors determined the target perceived color; that is, orange appeared reddish in the yellow corners/red background display and was perceived as yellowish in the opposite configuration. In the control conditions we obtained simultaneous color contrast with the background (see Fig. III.43-5).

We also tested whether lightness induction produced by perceptual grouping depends on the physical contrast between induced and inducing elements (Galmonte, Soranzo, & Agostini, 2006). In order to test this hypothesis, we systematically manipulated the relative physical

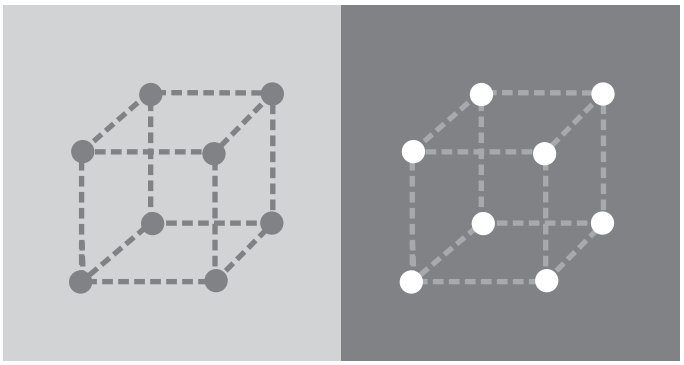


Figure III.43-4. When disks having the same area and luminance replace the inducer corners, the induction effect goes in the direction of local contrast (Image by Revised from [Agostini, T., & Galmonte, A. (2002). Perceptual organization overcomes the effect of local surround in determining simultaneous lightness contrast. *Psychological Science, 13*(1), 88–92]).

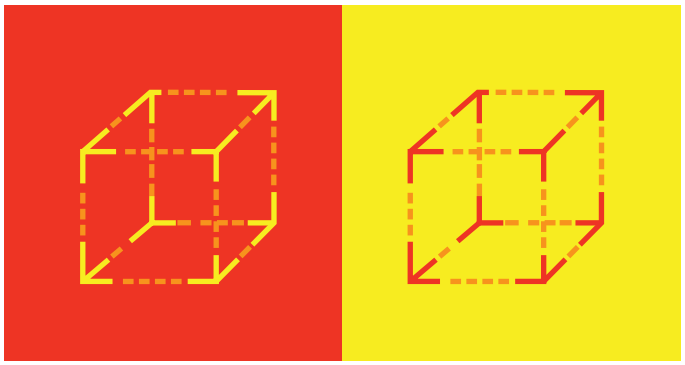


Figure III.43-5. The dashed lines belonging to the red inducer corners appear more yellowish than the dashed lines belonging to the yellow inducer corners. (Image by Alessandra Galmonte and Tiziano Agostini.)

contrast among the regions forming the original Agostini and Galmonte configuration (see Fig. III.43-6). Observers were asked to match the lightness of both the inducing and the induced element on a Munsell scale placed on the same background of the experimental display. We found that (a) global induction overcame local induction for all the tested physical contrasts, (b) global induction was stronger for decrements, (c) the amount of the global induction did not depend on the size of physical contrast, and (d) global induction seemed to be modulated by the physical contrast of increments only.

Another study was designed to test whether the lightness induction produced by perceptual grouping depends on the configuration spatial articulation (Galmonte, Gherzil, & Agostini, 2008). In order to test this hypothesis, we systematically manipulated the spatial articulation, starting from the original Agostini and Galmonte configuration and then gradually decreasing its complexity (see Fig. III.43-7). We found that (a) global induction overcame local induction for all the tested displays, (b) global induction seemed to be stronger for decrements, (c) global induction seemed to be modulated by the spatial articulation level for increments only, and (d) global inductions seemed to be symmetrically distributed for three-dimensional displays only.

DISCUSSION OF RESEARCH AND THEORIES

Simultaneous contrast is usually explained according to lateral inhibition among receptors. Recording from the nerve fibres of the eye of the *Limulus,* it has been found (Hartline, Wagner, & Ratliff, 1956) that the firing rate of a receptor stimulated by a certain amount of light is diminished when the neighboring receptors are stimulated by the same illumination. This physiological mechanism is called lateral inhibition because it is transmitted laterally across the retina. In humans, horizontal and amacrine cells of the retina can provide a pathway for transmitting inhibitory signals from one receptor to another.

Since the milestone works by Cornsweet (1970) and Hurvich and Jameson (1957), lateral inhibition models have been considered the main explanation of simultaneous contrast effects. These authors claim that the receptors stimulated by the light background send inhibition to the

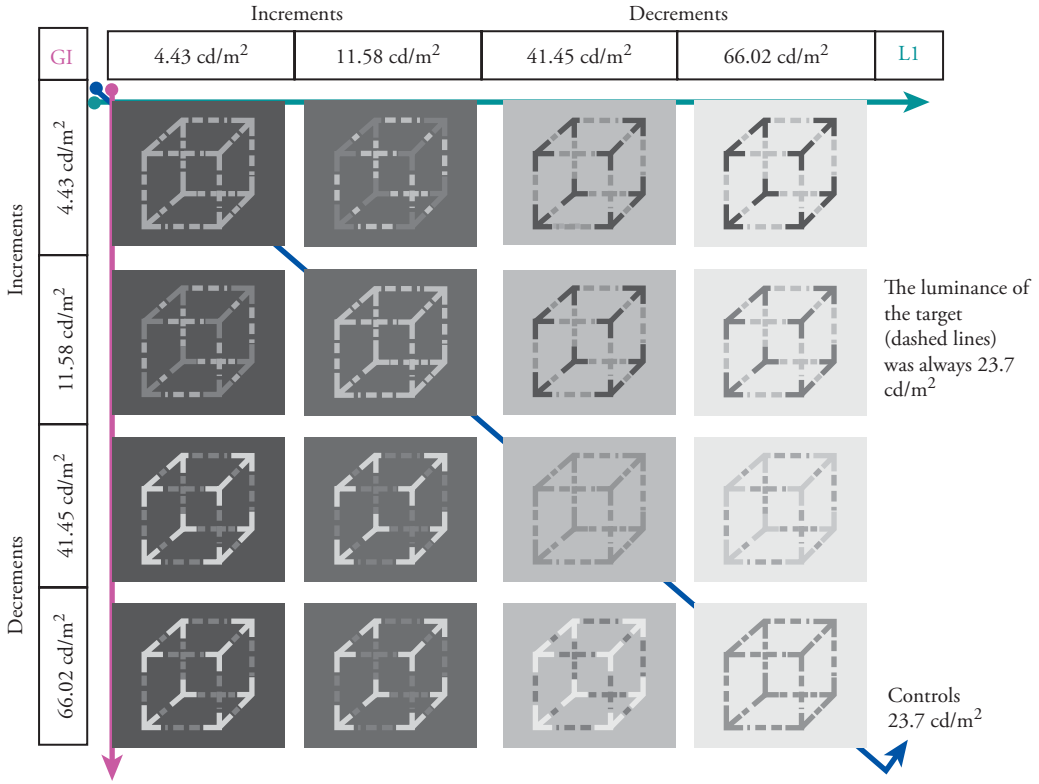


Figure III.43-6. The experimental stimuli: 16 reversed lightness contrast Necker cubes resulting from the combination of two variables. GI = global inducers; LI = local inducers. (Image by Alessandra Galmonte and Tiziano Agostini.)

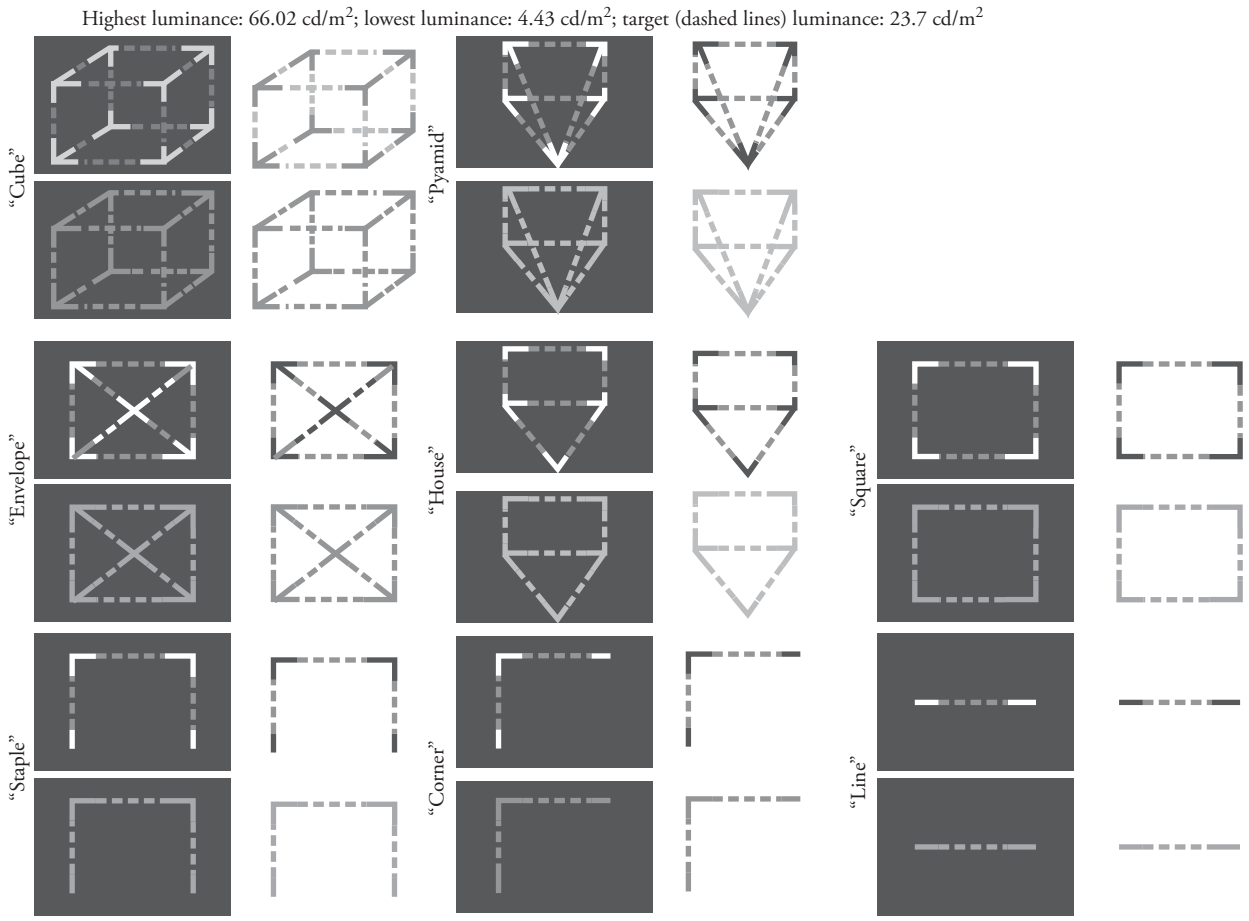


Figure III.43-7. The experimental stimuli in which the degree of spatial articulation has been manipulated. (Image by Alessandra Galmonte and Tiziano Agostini).

receptors stimulated by the grey surface, causing its perceptual darkening. However, the receptors stimulated by the dark background send little inhibition, and, therefore, there is not a darkening effect.

Lateral inhibition is the predominant explanation for contrast phenomena. However, other examples of simultaneous contrast have been offered where higher-level processes must be taken into account (Adelson, 1993; Agostini & Proffitt, 1993; Benary, 1924; Gilchrist, 1980; Logvinenko, 1999; Schirillo, 1999a, 1999b; Todorović, 1997; White, 1979).

Contrast phenomena have also been explained according to principles of perceptual organization (Gilchrist et al., 1999; Wertheimer, 1923); that is, the perceived color of a surface is mainly determined by the global contrast with the other colored surfaces to which it perceptually belongs and not simply by low-level processes of local contrast involving lateral inhibition.

The concept of perceptual belongingness was introduced by Benary in 1924, who, starting from a Wertheimer (1923) observation, stated that lightness perception was not a mere result of local interactions among visual features but also depended on their perceptual organization. This concept has guided a number of scientists in the investigation of the principles of visual processing. For example, it led Musatti (1953) to create the theory of perceptual equalization (*eguagliamento*). In Kanizsa's (1979) work, belongingness played a central role in explaining a variety of simultaneous contrast configurations. Later, Agostini and Proffitt (1993) showed that contrast effects depend on belongingness induced by common fate, influencing the theory by Gilchrist et al. (1999). Nevertheless, for a long time, this concept has been considered blasphemous regarding the processes underlying contrast, while at the same time a similar concept emerged within the feature-binding theory of Treisman and Gelade (1980) and experienced great success in explaining the mechanisms underlying the perception of visual objects. As a consequence, this almost 90-year-old concept provides a much-needed framework to test the theory that synchronization of neuronal discharges underlies the interaction between visual features (Singer, 1999), and it enables the investigation of the causal relationships between the synchronization as a physiological counterpart of belongingness (or binding) on one hand and the changes in the neuronal discharge rates evoked by lightness contrast on the other.

One important issue for which the interaction between contrast and belongingness becomes important is the understanding of the neurophysiological mechanisms underlying the integration of features, or the neurophysiological mechanisms underlying grouping, binding, and belongingness. The so-called binding by synchrony hypothesis was proposed earlier in purely theoretical formulations by Milner (1974), Grossberg (1976), and von der Malsburg

(1981/1994), but the Singer lab was the first to obtain experimental evidence supporting the potential role of synchrony as a relational code. Synchronization probability appeared to depend on the Gestalt criteria that the visual system applies in order to accomplish scene segmentation and perceptual grouping (Gray, König, Engel, & Singer, 1989; Gray & Singer, 1989). According to the binding by synchrony hypothesis, the neurons that respond to the features of the same perceptual object synchronize their action potentials but do not synchronize with the neurons responding to other objects in the visual scene. Experimental findings with respect to the changes of the strength of synchronization support this hypothesis, but there is still an important problem. It is difficult to prove that these changes in synchronization have functional significance. In other words, these changes could be a side effect (or an epi-phenomenon) of some other unknown process. Therefore the changes in the strength of synchronization in visual cortex must play a functional role in cortical computations.

The interaction between belongingness and contrast that Benary (1924) discovered provides the necessary mechanism of interaction between the putative mechanisms responsible for perceptual organization and the commonly studied brain mechanisms that are based on the firing rate responses of cortical neurons. The vigorousness of neuronal firing rates is highly correlated with, and thus believed to underlie, our perception of lightness and contrast. For example, when a stimulus is placed over a neuron's receptive field, the neuron produces strong responses. When an additional stimulus is placed in the surround, the response is inhibited. It is believed that due to this inhibition we perceive lower luminance of the stimulus. This inhibition is dependent on perceptual organization. For example, a surround stimulus that has a different orientation than the center stimulus produces much less inhibition. Thus the surround stimulus that is much less likely to be perceived as belonging to the center produces much less inhibition. Biederlack et al. (2006) have been able to show that this is indeed the case. When surround stimulus is placed at a larger distance from the receptive field (greater than about 4° of visual angle), surround stimuli can efficiently inhibit the responses of neurons only if the synchronization between neurons is strong. Inhibition is produced only by collinear surround stimuli and only if synchronization is strong. This experiment, which combines the principles of perceptual organization and the contrast between different parts of the stimulus, investigates the interaction between the neurophysiological counterparts of two different perceptual phenomena: belongingness and contrast.

We propose that the term "perceptual belongingness" refers to the result of a connection process unifying the specific perceptual attributes arising from spatial articulation (Agostini & Galmonte, 1999). We use the term "perceptual spatial articulation" to refer to the result of a perceptual organization process of selecting and highlighting specific physical spatial relationships of a visual pattern. Perceptual belongingness, however, refers to the result of a connection process that unifies the specific perceptual attributes arising from spatial articulation.

Consider the following example. A number of identical small disks are equidistantly spaced along the perimeter of a circle. Perceptual organization processes acting on the physical attributes of this pattern will give rise to perception of a circle of small disks. The circle arises from perceptual spatial articulation. Each individual disk appears to belong to the circle. It is in this way that perceptual spatial articulation determines the specific belongingness relationships among the elements.

We propose that perceptual spatial articulation also determines the perception of the surface colors of areas in the visual field. Perceptual spatial articulation is an operational concept, helping to formulate predictions about the lightness value taken by a particular region of the visual field. We assume that spatial articulation is low when there are few different reflectances and simple spatial relationships and that it becomes greater when there are more areas of different reflectance and/or their spatial relationships are more complex. Low spatial articulation determines univocal belongingness relationships, while high spatial articulation produces manifold belongingness relationships.

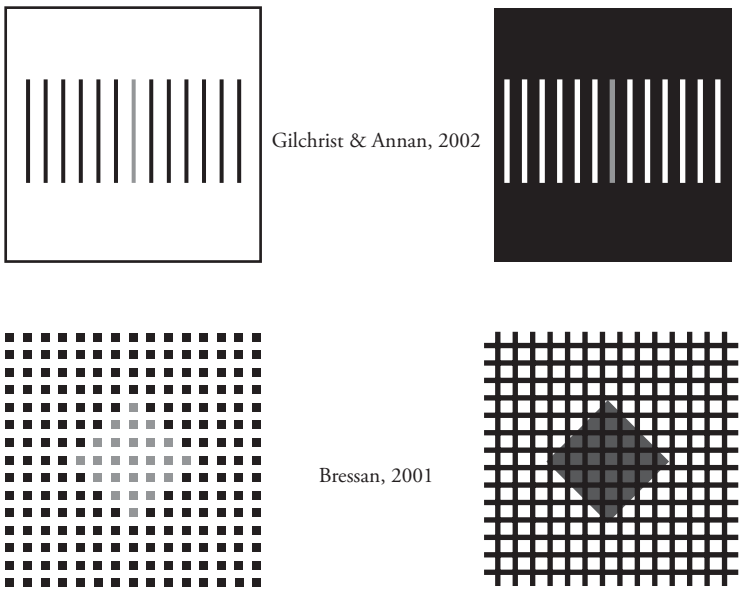


Figure III.43-8. "Reversed contrast" displays by Gilchrist and Annan (2002) and Bressan (2001) (Images by Revised from [Gilchrist, A. L., & Annan, V. (2002). Articulation effects in lightness: Historical background and theoretical implications. Perception, 31, 141–150] and [Bressan, P. (2001). Explaining lightness illusions. Perception, 30, 1031–1046]).

Therefore, the lightness of a region is the result of the influence of one or more belongingness relationships.

Two other configurations in the literature have been considered cases of reversed contrast (Bressan, 2001; Gilchrist & Annan, 2002; see Fig. III.43-8). The attempt made by these authors to find a simpler configuration failed to identify a locus in which perceptual belongingness and local surround can compete independently from other factors. In fact, it has been demonstrated that the induction effect observed in these configurations is due to an assimilation and not to a reversed contrast phenomenon—see Figure III.43-9 for a pictorial demonstration (Agostini & Galmonte, 2001; Agostini, Murgia, & Galmonte, 2014). The risk of studying perceptual belongingness, which is a high-level process, by using oversimplified stimuli is that it may introduce low-level confounding variables (see Fig. III.43-10 for a pictorial explanation).

From a strictly methodological point of view, the debate about the use of the theoretical concept of "perceptual belongingness," which is difficult to operationalize, is still open. The progress of the results in a field of research is closely linked to its theoretical assumptions, and from this viewpoint the concept of perceptual belongingness has been very fruitful when combined with experimental protocols. Therefore it seems reasonable to ask ourselves whether it

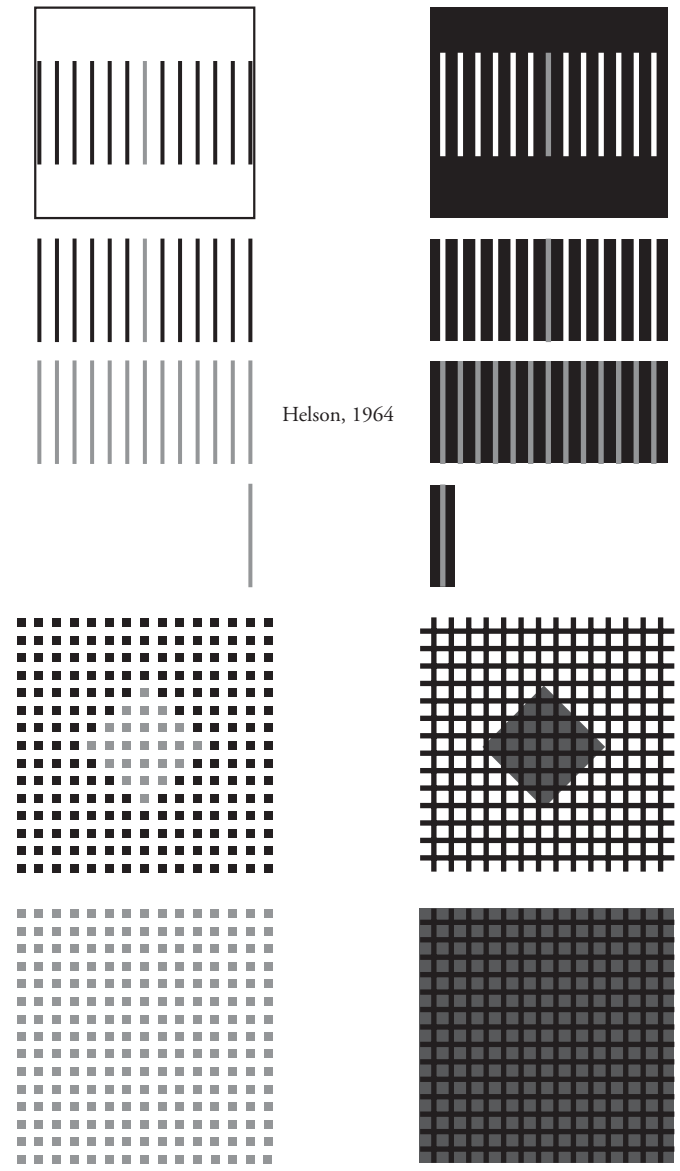


Figure III.43-9. Pictorial demonstration that both Annan and Gilchrist (2002) and Bressan (2001) configurations can be interpreted as assimilation effects rather than reversed contrast ones. In fact, by substituting all the inducing elements that are supposed to elicit contrast by grouping with elements having the same luminance of the target(s), the outcome is the same as that used by Helson (1964) to demonstrate assimilation effects (Image by Revised from [Agostini, T., Murgia, M., & Galmonte, A. (2014). Reversing the reversed contrast. Perception, 43, 207–213]).

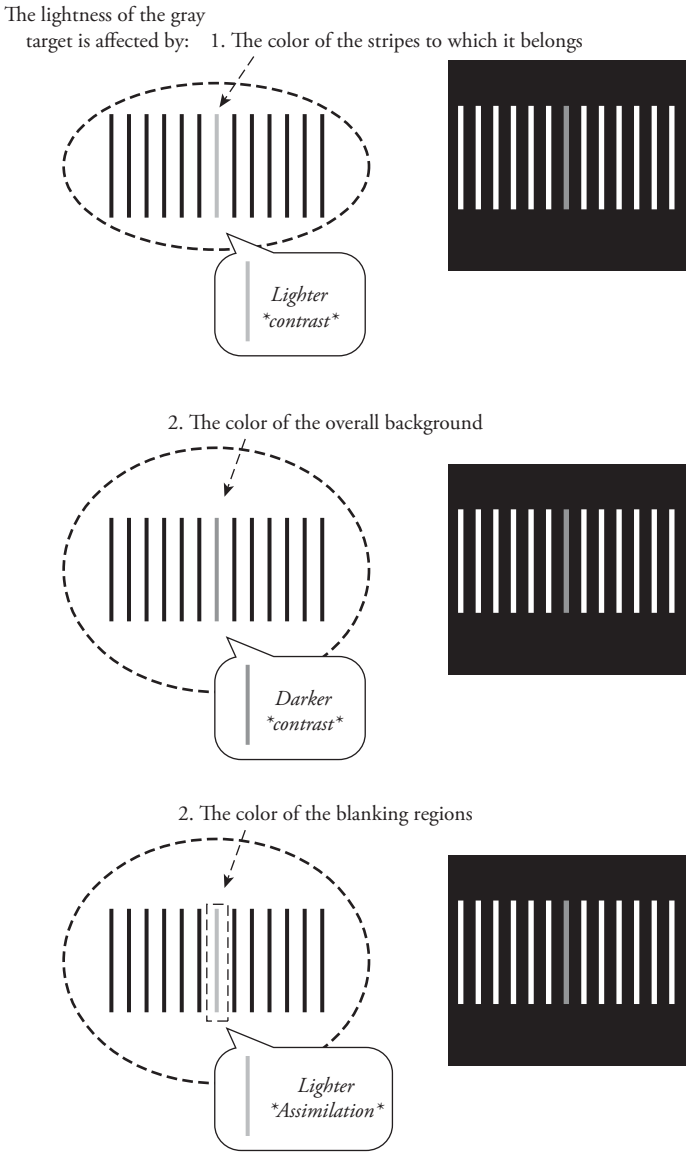


Figure III.43-10. The ambiguity of the Annan and Gilchrist (2002) "reversed contrast" display. In this case, the grey target region is affected by three different factors: (a) contrast with the belonging elements (i.e. a high-level factor), (b) contrast by the overall surround (i.e., a low-level factor), and (c) assimilation by the flanking regions (i.e., a low-level factor). This means that, using oversimplified stimuli to study perceptual belongingness, which is a high-level process, can introduce low-level confounding variables (Image by Revised from [Agostini, T., Murgia, M., & Galmonte, A. (2014). Reversing the reversed contrast. Perception, 43, 207–213]).

is wise to give up this concept to look for other ones that are easier to operationalize but probably are less fertile in relation to the potential results. As James (1890/1950) stated in his *Principles of Psychology*:

> The boundary-line of the mental is certainly vague. It is better not to be pedantic, but to let the science be as vague as its subject . . . we gain much more by a broad than by a narrow conception of our subject. At a certain stage in the development of every science a degree of vagueness is what best consists with fertility. (p. 6)

REFERENCES

Adelson, E. H. (1993). Perceptual organization and the judgment of brightness. *Science, 262*, 2042–2044.

Agostini, T., Da Pos, O., Galmonte, A., & Mauri, A. (2009). Does colour undergo the same reversed contrast effect observed for lightness in Agostini and Galmonte's Necker cube display? *Perception, 38*(ECVP Abstract Suppl.), 109.

Agostini, T., & Galmonte, A. (1999). Spatial articulation affects lightness. *Perception & Psychophysics, 61*(7), 1345–1355.

Agostini, T., & Galmonte, A. (2001). Contrast or assimilation? Two ambiguous cases. *Perception, 30*(ECVP Abstract Suppl.), 7.

Agostini, T., & Galmonte, A. (2002). Perceptual organization overcomes the effect of local surround in determining simultaneous lightness contrast. *Psychological Science, 13*(1), 88–92.

Agostini, T., Murgia, M., & Galmonte, A. (2014). Reversing the reversed contrast. *Perception, 43*, 207–213.

Agostini, T., & Proffitt, D. R. (1993). Perceptual organization evokes simultaneous lightness contrast. *Perception, 22*, 263–272.

Benary, W. (1924). Beobachtungen zu einem experiment über helligkeitskontrast. *Psychologische Forschung*, *5*, 131–142. (English translation: The influence of form on brightness contrast. In W. D. Ellis [Ed.], *A source book of Gestalt psychology* [pp. 104–109]. London: Routledge & Kegan Paul, 1939.)

Biederlack, J., Castelo-Branco, M., Neuenschwander, S., Wheeler, D. W., Singer W., & Nikolić, D. (2006). Brightness induction: Rate enhancement and neuronal synchronization as complementary codes. *Neuron, 52*, 1073–1083.

Bressan, P. (2001). Explaining lightness illusions. *Perception, 30*, 1031–1046.

Cornsweet, T. N. (1970). *Visual perception*. New York: Academic Press.

Galmonte, A., Gherzil, A., & Agostini, T. (2008). The role of spatial articulation in Agostini & Galmonte reversed lightness induction Necker cube. *Perception, 37*(Suppl.), 145.

Galmonte, A., Soranzo, A., & Agostini, T. (2006). The role of physical contrast in Agostini & Galmonte's reversed lightness-induction Necker cube. *Perception, 35*(ECVP Abstract Suppl.), 180.

Gilchrist, A. L. (1980). When does perceived lightness depend on perceived spatial arrangement? *Perception & Psychophysics, 28*, 527–538.

Gilchrist, A. L., & Annan, V. (2002). Articulation effects in lightness: Historical background and theoretical implications. *Perception, 31*, 141–150.

Gilchrist, A. L., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., . . . Economou, E. (1999). An anchoring theory of lightness perception. *Psychological Review, 106*(4), 795–834.

Gray, C. M., König, P., Engel, A. K., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature, 338*, 334–337.

Gray, C. M., & Singer, W. (1989). Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proceedings of the National Academy of Sciences USA, 86*, 1698–1702.

Grossberg, S. (1976). Adaptive pattern classification and universal recoding: II. Feedback, expectations, olfaction, and illusions. *Biological Cybernetics, 23*, 187–202.

Hartline, H. K., Wagner, H. G., & Ratliff, F. (1956). Inhibition in the eye of Limulus. *Journal of General Physiology, 39*, 651–673.

Helson, H. (1964). *Adaptation-level theory*. New York: Harper & Row.

Hurvich, L. M., & Jameson, D. (1957). An opponent process theory of color vision. *Psychological Review, 64*, 384–404.

James, W. (1950). *Principles of psychology*. New York: Holt. (Original work published 1890)

Kanizsa, G. (1979). *Organization in vision: Essays on Gestalt perception*. New York: Praeger/Greenwood.

Logvinenko, A. D. (1999). Lightness induction revisited. *Perception, 28*(7), 803–816.

Milner, P. M. (1974). A model for visual shape recognition. *Psychological Review, 81*(6), 521–535.

Musatti, C. (1953). Luce e colore nei fenomeni del contrasto simultaneo, della costanza e dell'eguagliamento. *Archivio di Psicologia, Neurologia e Psichiatria, 5*, 544–577.

Schirillo, J. A. (1999a). Surround articulation: I. Brightness judgment. *Journal of the Optical Society of America, 16*, 793–803.

Schirillo, J. A. (1999b). Surround articulation: II. Lightness judgments. *Journal of the Optical Society of America, 16*, 804–811.

Singer, W (1999). Neuronal synchrony: A versatile code for the definition of relations? *Neuron, 24*, 49–65.

Todorović, D. (1997). Lightness and junctions. *Perception, 26*(4), 379–394.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 16*, 97–134.

von der Malsburg, C. (1994). The correlation theory of brain function. In E. Domany, J. L. van Hemmen, & K. Schulten (Eds.), *Models of neural networks II* (pp. 95–120). Berlin: Springer. (Original work published 1981)

Wertheimer, M. (1923). Untersuchungen zur lehre von der Gestalt. *Psychologische Forschung*, *4*, 301–350. (English translation: Laws of organization in perceptual forms. In W. D. Ellis (Ed.), *A source book of Gestalt psychology*. London: Routledge & Kegan Paul, 1939].

White, M. (1979). A new effect of pattern on perceived lightness. *Perception, 8*, 413–416.

Chapter 44
Changing the Chevreul Illusion by a Background Luminance Ramp

János Geier and Mariann Hudák

PRESENTATION AND DESCRIPTION OF THE EFFECT

The classical Chevreul illusion comprises spatially uniform grey bands of different luminance that seem inhomogeneous, as if they were crimped: each band looks darker on one side and lighter on the other (Fig. III.44-1). According to Ratliff (1965), this illusion is attributed to Michel Eugène Chevreul (1786–1889), who placed spatially uniform bands of gradually increasing luminance next to each other, whereby he discovered the illusion.[1]

We placed the Chevreul staircase in a luminance ramp background, which considerably affected the illusion: it significantly increases or decreases, depending on the progression of the ramp relative to the staircase. When the progression of the staircase is identical with that of the ramp, the illusion is enhanced, whereas when the staircase and the ramp progress in opposite directions, the illusion ceases. (We placed the staircase into the ramp so that the staircase itself remained physically unchanged.)

This phenomenon can directly be experienced by looking at Figure III.44-2, in which we placed two physically identical staircases of opposite progressions in a luminance ramp background. Note that the change in the illusory effect is equally strong through the entire height of each step. In other words, the change of the illusion is not limited to the immediate neighborhood of the upper and lower edges of the steps, where they adjoin the ramp. This is so even when any inner point of the staircase far from the ramp is fixated for several seconds.

WHY THE EFFECT IS INTERESTING AND RELEVANT FOR VISION RESEARCH

The significance of the ramp effect is that it challenges the generally accepted explanation of the Chevreul illusion. Traditionally, the Chevreul illusion has been explained in terms of lateral inhibition, which means that lighter areas projected onto the retina inhibit the sensitivity of neighboring retinal areas. In line with this classical principle, the reason for the perceived inhomogeneity in the physically homogeneous steps is that the side of each step neighboring a lighter one receives more inhibition than its other side (Ratliff, 1965). In Figure III.44-1, the equivalent textbook explanation (e.g., Blake & Sekuler, 2006; Goldstein, 2009) is illustrated, according to which the proportion of lateral inhibition received by the on-center/off-surround (i.e., antagonistic) receptive fields located at the two opposite sides of each step is different. This difference in the relative amount of lateral inhibition is assumed to cause the illusory inhomogenity within each step.

The principle of lateral inhibition stems back as early as the 19th century, when it seemed to explain many of

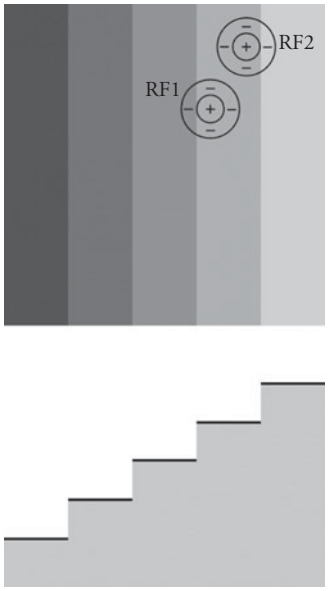


Figure III.44-1. The classical Chevreul illusion and its classical explanation. The physical luminance cross-section of the midline of the staircase is displayed in the bottom part of the figure. The steps adjacent to each other are physically homogeneous; however, they seem inhomogeneous (crimped). The side of each step adjoining a lighter step seems darker than its other side. The classical explanation is demonstrated by the on-center/off-surround receptive fields RF1 and RF2. The left side of the RF1 surround reaches into a darker region, while the right side of RF2 surround covers a lighter area; therefore, the proportion of lateral inhibition received by RF2 is larger than that of RF1. According to the classical explanation, this difference in the proportion of inhibition causes the left side of each step to be lighter than the right side. (Geier, J., & Hudák M. (2011). Changing the Chevreul illusion by a background luminance ramp: Lateral inhibition fails at its traditional stronghold—A psychophysical refutation. *PLoS ONE, 6*(10), e26062.)

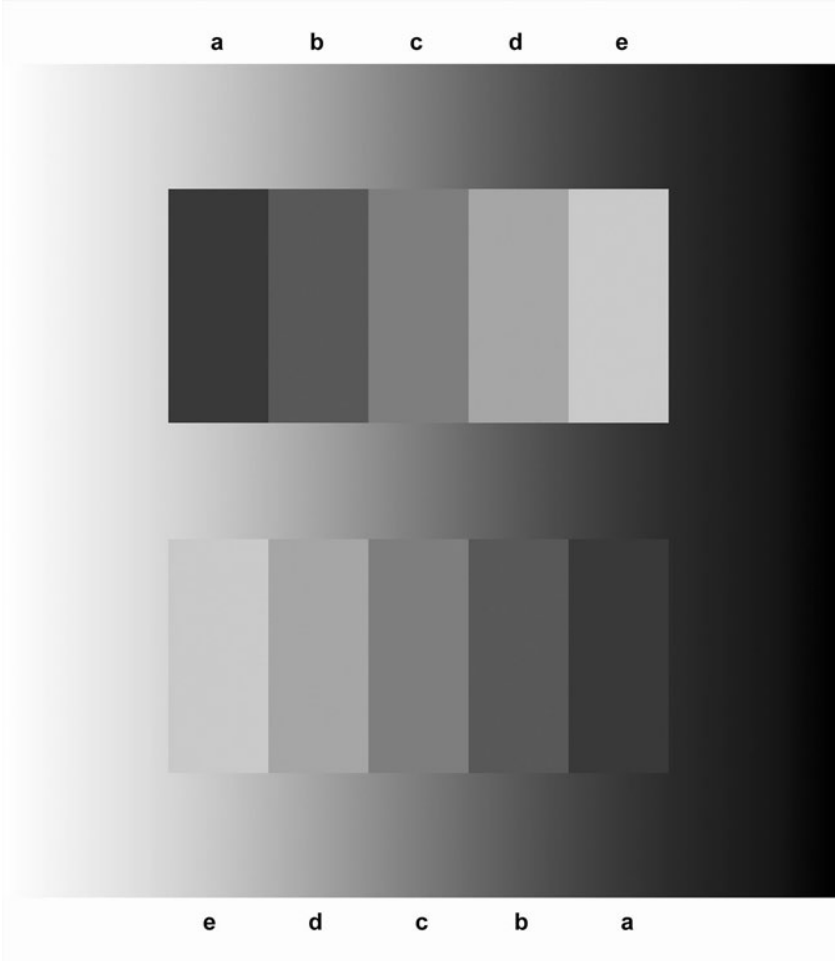


Figure III.44-2. The effect of the luminance ramp background. Two physically identical Chevreul staircases of opposite progressions were placed in a luminance ramp background. (Identical letters indicate the steps of physically identical luminance.) It can be seen that, due to the ramp, the illusion has significantly changed: the illusion ceases if the progression of the staircase is opposite to that of the ramp (upper staircase), while it is strongly enhanced when the progressions of the ramp and that of the staircase are identical (lower staircase). Note that the illusion has changed through the entire height of the steps. (Geier, J., & Hudák M. (2011). Changing the Chevreul illusion by a background luminance ramp: Lateral inhibition fails at its traditional stronghold—A psychophysical refutation. *PLoS ONE, 6*(10), e26062.)

the then-known brightness illusions, and it has been considered the basic mechanism of early vision (Ratliff, 1965). Antagonistic circular receptive fields implementing lateral inhibition in the retina (Kuffler, 1953) are described mathematically by the difference of Gaussians (DoG) model (Rodieck, 1965). Several current multiscale spatial filtering models of brightness perception also build upon the DoG model with more or less supplementation, retaining its local nature (e.g., Blakeslee & McCourt, 1999; Kingdom & Moulden, 1992). All these models vary the DoG principle so that they either use series of DoG filters of various spatial frequencies or their variants with an elongated shape (ODOG). In line with the terminology found in the literature, we hereafter identify the concept of lateral inhibition with models using DoG-like filters, including multiscale models and models using elongated filters as well as any qualitative explanations referring to such (e.g. the classical textbook explanation).

The aim of the DoG model (as well as other models of brightness perception) is to predict the brightness (perceived luminance) distribution from the physical luminance distribution of an image. The input of such a model is an image corresponding to the physical luminance distribution, while another image is expected as output, in which the intensities correspond to human perception.

In our modified Chevreul illusions, the replacement of the original white background with a luminance ramp background causes physical luminance chance exclusively *outside* the area of the staircase, while no physical change has occurred within the staircase (cf. the modification by Morrone, Burr, & Ross, 1994, who changed the illusion by placing lines *within* the steps). Classical lateral inhibition-based explanations, as pointed out already, build exclusively upon luminance relations of the steps *within* the staircase. Therefore, if classical lateral inhibition–based explanations were tenable for the Chevreul illusion, then the perception within the steps should not be changed by the ramp. This explanation, at the best, can predict change merely near the upper and lower boundaries of the staircase but not through the entire height of the steps. This is in contradiction with the phenomenon that the illusion has changed through the entire vertical height of the staircase merely due to the surrounding luminance ramp. Consequently, it can be concluded that the classical explanation based on lateral inhibition is not tenable for the ramp effect.

VARIANTS AND RELEVANT PARAMETERS FOR THE EFFECT

If we aim to find a new, unified explanatory principle for all these phenomena, we have to notice that placing the ramp around the staircase physically changes not only the area outside the staircase but the boundary edges of the staircase too. To decide which of these plays a more important role in the change of the Chevreul illusion, we placed

another, narrow ramp around the staircase, whose direction was opposite to that of the original, larger ramp.

The result of this modification involving a double luminance ramp can be directly observed in Figure III.44-3. It can be seen that although the area of the inner ramp is significantly smaller than that of the outer ramp, the inner one still governs the change in the Chevreul illusion. If the inner ramp is replaced by a homogeneous rectangle, then the two Chevreul staircases will be perceived as identical (progressing in opposite directions), and the outer ramp will have no effect.

This result supports that the upper and lower boundary edges of the staircase control the perceptual experience and not the area size of the ramp, since such a narrow ramp as half a degree can prevail against the effect of the much larger outer ramp. Therefore we conclude that it is the boundary edges in the image that govern perceptual experience instead of the large background areas, and long-range interactions should be supposed between edges and the areas enclosed by them.

DISCUSSION OF RESEARCH AND THEORIES

We summarize the description of these perceptual phenomena as follows (for detailed experimental results, see Geier & Hudák, 2011).

Result 1: In a Chevreul staircase with a homogeneous background, most observers (~ 70%) see the steps as crimped.

Result 2: On placing the staircase in a luminance ramp background of opposite progression, the illusion ceases, while on placing it into a ramp of identical progression, the illusion is significantly enhanced. (This was the case for 100% of our observers.)

Result 3: When the staircase is placed in a double luminance ramp, the inner one governs the perceptual experience even when its area is rather small compared to the outer one, and Results 1 and 2 also hold here for the perceptual experience.

Result 4: Regardless of the variant of the Chevreul staircase being observed (either the classical one with a homogeneous background or the single or double-ramped versions enhancing or ceasing the effect), the extent of the perceived homogeneity (or the inhomogeneity) of each step is equal within the entire height of the staircase. The illusion is of the same magnitude near the upper or the lower boundaries, as well as in the midline of the staircase.

The dynamic version of Figure III.44-3 is shown in video III.44-1 on the Companion website, which demonstrates the experimental task.

WHY IS LATERAL INHIBITION INSUFFICIENT HERE?

The main point of our criticism, as mentioned earlier, is that the classical lateral inhibition account of the Chevreul illusion considers merely the neighboring steps as the local surround of each step, and thus it cannot take the effect of

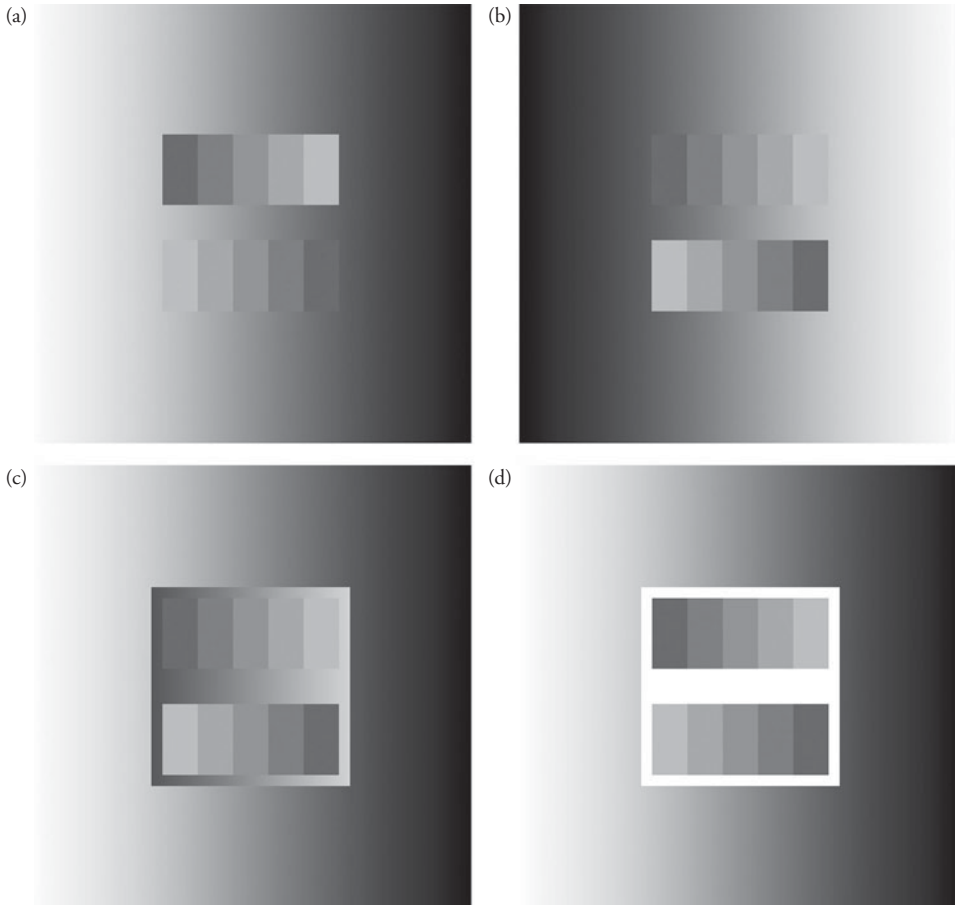


Figure III.44-3. The effect of a double-ramp background. The staircase pairs in the four images are physically identical; the upper and lower staircases in each image are also identical except for their progressions. On comparing (a) and (b), we can see that a ramp of opposite progression causes opposite effects. On comparing (c) with (a) and (b), we can see that the illusion in (c) is identical to that in (b). This is so, although the large outer ramp in (c) is identical to the one in (a). Therefore the small inner ramp dominates perception, whose direction is identical to that in (b). Finally, the upper and lower staircases in (d) look identical (except for their progressions); therefore the inner small area, the homogeneous white rectangle, is what dominates here too. (Geier, J., & Hudák M. (2011). Changing the Chevreul illusion by a background luminance ramp: Lateral inhibition fails at its traditional stronghold—A psychophysical refutation. *PLoS ONE, 6*(10), e26062.)

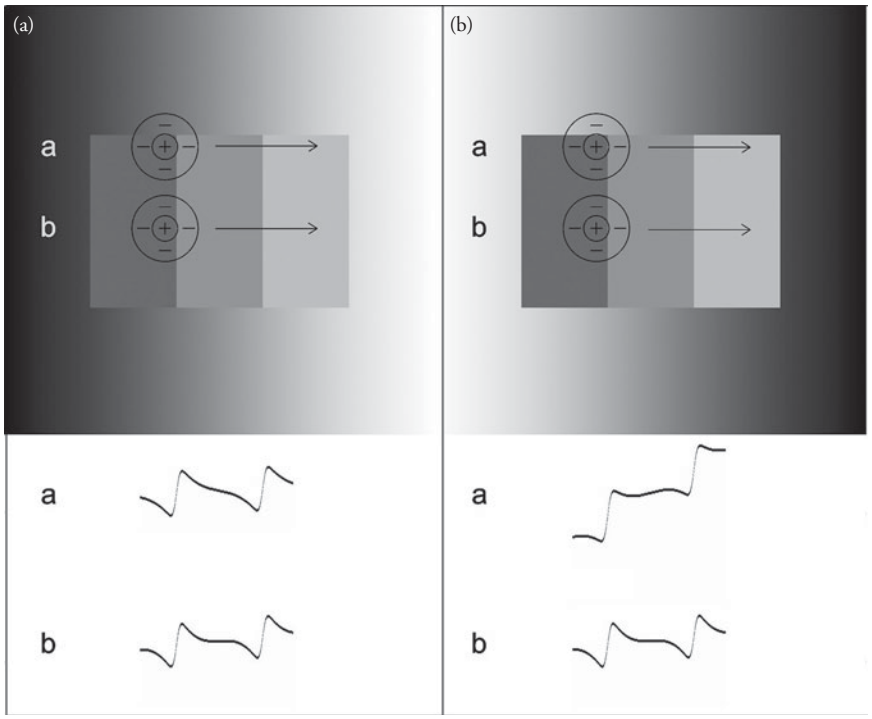


Figure III.44-4. The output of the DoG model for steps in ramps of opposite progressions. The two staircases in (a) and (b) are physically identical; only the progressions of the luminance ramp backgrounds differ. The ramp effect can be seen in spite of the explanatory signs drawn in the image. If the DoG filters are moved along the horizontal direction, as shown by the arrows, they will predict the brightness values shown in the brightness cross-section diagrams (graph a) and (graph b) below the image. The luminance cross-sections produced by the simulation of the DoG filter at cross-sections (graph a) and (graph b) are shown below the steps, respectively. On the one hand, the prediction of DoG filter (graph a) is somewhat similar to human perception, since it predicts a steeper slope in (a). On the other hand, though the steps in (b) are seen as totally flat, DoG filter (graph a) still predicts scalloping there. In addition, in the midline of the two staircases, no difference is predicted between (a) and (b) by DoG filter (graph b), contradictory to human perception, according to which the steps in (a) and (b) look largely different. Moreover, the predictions of (graph a) and (graph b) within each staircase show different brightness cross-sections, although the illusion is equally strong through the entire height of the staircase. (The cross-section diagrams were produced by the computer simulation of the DoG filter.) (Geier, J., & Hudák M. [2011]. Changing the Chevreul illusion by a background luminance ramp: Lateral inhibition fails at its traditional stronghold—A psychophysical refutation. *PLoS ONE, 6*(10), e26062.)

the ramp outside the staircase into account. We next analyze this in more detail.

DoG filters, corresponding to the classical explanation, are illustrated in the inner area and near the upper boundary edges of the staircases in Figure III.44-4. By comparing the cross-section diagrams predicted by DoG filters, numerous contradictions can be found with human perception. It can be seen that the DoG model predicts ramp effect only near the upper and lower boundary edges of the staircase, but no ramp effect is predicted in the middle areas of the steps. This contradicts both Result 2, according to which ramps of opposite progressions cause opposite effects on the illusion, and Result 4, according to which the change of the illusion is equally strong through the entire height of the staircase. These contradictions are not surprising, since DoGs located in the inner parts of the staircase do not extend outside the staircase boundaries; therefore, they are not influenced by the ramp.

Another side effect of such a simple DoG filtering is the blur of the step edges, as can be seen in the cross-section diagrams. Multiscale models attempt to handle this problem by applying DoG filters of small diameters to avoid blurring, as well as very large ones to ensure that remote points can influence inner parts of large homogeneous areas. Therefore it could be reasonable to think that multiscale models can predict the phenomena presented in our images. However, such models fail to predict our double-ramped variants. The reason for this is that the large filters (e.g. in the ODOG model, the largest filter diameter is 36° including the surround) are influenced by the outer ramp to such an extent that the outer ramp will dominate the predicted perceptual experience, contradictory to human perception (for a more detailed analysis, see Geier & Hudák, 2011). In case of the application of small filters only, however, the inner parts of the staircase will not be influenced even by the inner ramp, as shown already in Figure III.44-4.

Consequently, DoG-based models fail to predict the ramp effect phenomena, since neither small nor large filters are able to capture these changes, irrespective of whether they are circularly symmetric or elongated.

CONCLUSIONS

On the basis of our results, our conclusions are the following.

Conclusion 1: Lateral inhibition-based models cannot capture the ramp effects presented here.

Conclusion 2: It is the edges that play the most significant role in the change of the illusion.

Conclusion 3: The edges also obstruct effects coming from farther edges (here the outer edge of the inner ramp prevents the effects coming from the direction of the outer ramp from spreading into the staircase).

Conclusion 4: There is a long-range interaction between edges and areas enclosed by them.

These conclusions might extend beyond the Chevreul illusion embedded in background ramp(s). We regard these conclusions as generally valid to brightness perception but not limited to the brightness phenomena introduced here.

As it has been shown here in detail, DoG models fail to give a unified explanation for these phenomena. Such models are built upon the weighted sum (convolution) of areas covered by single DoG filters; therefore they are essentially sensitive to appropriately weighted average intensities of larger or smaller portions of the image. The accentuated role of edges in the generation of the illusion is not included in DoG models, nor is their segmenting role. Finally, DoG models do not apply any interaction between filters remotely located from each other.

By means of the foregoing, we proved directly "only" that the ramped versions of the Chevreul illusion cannot be accounted for by the DoG model. It could be argued that the DoG model is still suitable to explain the classical Chevreul illusion presented on a white background. Nonetheless, let us consider the following: by the introduction of the variations with luminance ramp backgrounds, the classical white-backgrounded version has become merely a special case of the broader range of the Chevreul phenomena (i.e., here the slope of the background ramp is zero). It has been shown that the DoG model fails to provide a unified account for the ramped versions; therefore, a new model should be sought. To our knowledge, no such model exists at present in the literature. However, it is certain that once such a model is developed, it should obviously be able to capture both the ramped versions and the classical Chevreul illusion as well.

If the prediction of a model more or less agrees with the perceptual facts, it is useful. However, it is not sufficient in itself, since it might happen that this agreement is only apparent in a special case. What can be expected from a good model in principle is that it should capture the essence of the modeled process. This is the reason for developing models at all: to understand processes and phenomena better. The DoG model fails to capture the ramped versions; therefore, it is clear that it fails to capture the essence of these phenomena. Why would one think, therefore, that it captures the essence in the classical uniform-backgrounded case?

Therefore we base our claim that the lateral inhibition-based models are refuted by the ramped versions on the basis of this line of thought. The principle of lateral inhibition is unable to capture the Chevreul illusion since it fails to capture the essence of the broader range of phenomena of which the classical Chevreul illusion is a special case.

WHERE DO WE GO FROM HERE?

We are now in need of a model that can universally handle the points we claim in our conclusions. The most suitable candidates for this are the filling-in type of models. The prototype of such models (the "standard diffusion model") is what Cohen and Grossberg (1984) applied in one dimension, and, after them, Grossberg and Todorović (1988) extended it to two dimensions. This so-called CGT model was further developed by others, but as it turns out from Gilchrist's review (Gilchrist, 2006, pp. 206–207), its basic principle is practically unchanged even now.

The main point of the CGT model, in short, is that after the allocation of the edges, the areas enclosed by the edges are filled in by a diffusion process. At the same time, edges are also assigned an obstructive role. These principles are fully in line with our conclusions: it is the edges that govern the process; they also have an obstructive role, and the basis of the long-range interactions between the edges and the areas enclosed by them is the diffusion process.

However, Gilchrist's (2006) comment that the CGT model is unable to handle the staircase luminance pattern (i.e., the Chevreul illusion) should be taken into account. He mentions the following note by Pessoa, Mignolla, and Neumann (1994) on the CGT model: Perhaps an even greater challenge to filling-in models is a luminance staircase distribution. The "steps" of the staircase presumably block diffusion, and it is not evident how a filling-in model can predict that different steps appear with different brightnesses (since "border contrast" is the same everywhere). (p. 2202)

Therefore the CGT model needs an essential correction: the issue of the brightness of areas separated by edges should be solved. Hopefully this correction will sooner or later be achieved by someone.

ACKNOWLEDGMENTS

The contribution of MH to this chapter was supported by Grant TÁMOP–4.2.2.B-10/1--2010-0009.

This chapter is the shortened version of J. Geier & M. Hudák (2011), "Changing the Chevreul Illusion by a Background Luminance Ramp: Lateral Inhibition Fails at Its Traditional Stronghold—A Psychophysical Refutation," *PLoS ONE 6*(10), e26062.

NOTE

1. Many textbooks (e.g. Blake &, Sekuler, 2006; Goldstein, 2009) refer to the Chevreul illusion as Mach bands. However, according to Ratliff (1965, p. 41) and von Békésy (1967), Mach bands are seen when the linearity of a luminance ramp breaks.

REFERENCES

Blake, R., & Sekuler, R. (2006). *Perception*. Boston: McGraw-Hill.

Blakeslee, B., & McCourt, M. E. (1999). A multiscale spatial filtering account of the White effect, simultaneous brightness contrast and grating induction. *Vision Research*, *39*, 4361–4377.

Cohen, M. A., & Grossberg, S. (1984). Neural dynamics of brightness perception: Features, boundaries, diffusion, and resonance. *Perception & Psychophysics, 36*, 428–456.

Geier, J., & Hudák M. (2011). Changing the Chevreul illusion by a background luminance ramp: Lateral inhibition fails at its traditional stronghold—A psychophysical refutation. *PLoS ONE, 6*(10), e26062.

Gilchrist, A. (2006). *Seeing black and white*. Oxford: Oxford University Press.

Goldstein, E. B. (2009). *Sensation and perception*. Belmont, CA: Wadsworth Cengage Learning.

Grossberg, S., & Todorović, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena. *Perception & Psychophysics, 43*, 241–277.

Kingdom, F., & Moulden, G. (1992). A multi-channel approach to brightness coding. *Vision Research, 32*(8), 1565–1582.

Kuffler, S. W. (1953). Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology, 16*, 57–68.

Morrone, M. C., Burr, D. C., & Ross, J. (1994). Illusory brightness step in the Chevreul illusion. *Vision Research*, *34*(12), 1567–1574.

Pessoa L., Mignolla E., & Neumann H. (1994). A contrast- and luminance-driven multiscale network model of brightness perception. *Vision Research, 35*(15), 2201–2223.

Ratliff, F. (1965). *Mach bands: Quantitative studies on neural networks on the retina*. San Francisco, CA: Holden-Day.

Rodieck, R. W. (1965). Quantitative analysis of cat retinal ganglion cell response to visual stimuli. *Vision Research, 5*, 583–601.

von Békésy, G. (1967). *Sensory inhibition*. Princeton, NJ: Princeton University Press.

Chapter 45
The Curved Grid Non-Illusions

Eliminating Hermann's Spots and Lingelbach's Scintillation

János Geier and Mariann Hudák

PRESENTATION AND DESCRIPTION OF THE EFFECT

It may well be surprising at first sight to meet images in which no illusion can be seen in a compendium of visual illusions. Why are these "non-illusions" relevant here? We will understand the answer to this question if we take into consideration that these non-illusions are slight variants of the Hermann grid illusion known since the 19th century (Hermann, 1870) and that of Lingelbach's scintillating grid, which was invented about 20 years ago (Schrauf, Lingelbach, & Wist, 1997). Seeing no illusion in these curved grids violates our expectation, bearing in mind the classical explanation of the Hermann grid illusion.

The first illusion that we eliminated by curvature is the Hermann grid illusion. The original Hermann grid (Fig. III.45-1a) is a regular grid pattern comprising white horizontal and vertical stripes ("streets") on a black background. At the intersections of the streets, smoke-like, dark spots are perceived (except for those intersections that are being fixated). The illusion also works in inverted grids, when the background is white and the streets are black (Hering grid; see Wade, 2005). In this case, the illusion appears as cloud-like, white spots at the intersections. The illusion was first described by Ludimar Hermann in 1870; however, it is known that similar results had been reported earlier (for a review, see Spillmann, 1994, and Wade, 2005).

The second illusion that we eliminated by the same modification is the scintillating grid illusion. The recently known version of the scintillating grid was created by Schrauf et al. (1997). They set the lines of the original Hermann grid to grey, and they placed white disks at its intersections. As a result, a strong scintillation was perceived in these images. The first scintillating grid was introduced by Bergen (1985), who low-pass filtered the Hermann grid, which resulted in the perception of flashing at the intersections instead of static illusory spots. At this point, it could be suspected that the removal of high spatial frequencies was a necessary condition for scintillation to appear. However, Schrauf et al. elicited scintillation without low-pass filtering, that is, without removing high spatial frequencies. Thus they have proved that the phenomenon is not dependent on the removal of high spatial frequency; rather, the scintillation occurs when the intersections are lighter than the streets.

Our main results are illustrated in Figure III.45-1b and Figure III.45-2b. We have curved the streets of the grids in a sinus-like manner, which made both illusions disappear totally, while the grid character remained unchanged in both cases. It can be seen in Figure III.45-1b that the spots of the Hermann grid disappeared as an effect of the slight sine curvature. By applying the same distortion to the scintillating grid, the scintillation also stops (Fig. III.45-2b). It is noteworthy that for the illusion to disappear it is sufficient to raise the amplitude of the sine wave to less than 10% of its wavelength. Both of these phenomena were first introduced by J. Geier in a Hungarian visual symposium in 2001, which was followed by its presentation at ECVP 2004 in Budapest (Geier, Séra, & Bernáth, 2004).

Two videos (Video III.45-1, 45-2) of the sinusoid Hermann grid and scintillation grid is included on the Companion Website, in which the amplitude of the curves can be adjusted until the spots disappear or reappear.

WHY THE EFFECT IS INTERESTING AND RELEVANT FOR VISION RESEARCH

The transformation of two strong illusions to two non-illusions by a simple modification suggests at least two theoretical points. The first is that the classical lateral inhibition-based explanation of the Hermann grid illusion is not tenable. The second is that a common mechanism is to be found behind the two illusions, since the same modifications cease both effects.

Let us first analyze the explanation of the Hermann grid illusion. Hermann himself accounted for the phenomenon in accordance with the principle of lateral inhibition already in 1870 (Wade, 2005). The physiological realization of this principle was proposed by Baumgartner (1960), whose explanation was based on the center-surround antagonism of concentric receptive fields in the retina described earlier by Kuffler (1953).

Baumgartner's hypothesis (illustrated in Fig. III.45-1a) may be considered as the official textbook account of the Hermann grid illusion (Blake & Sekuler, 2006; Goldstein,

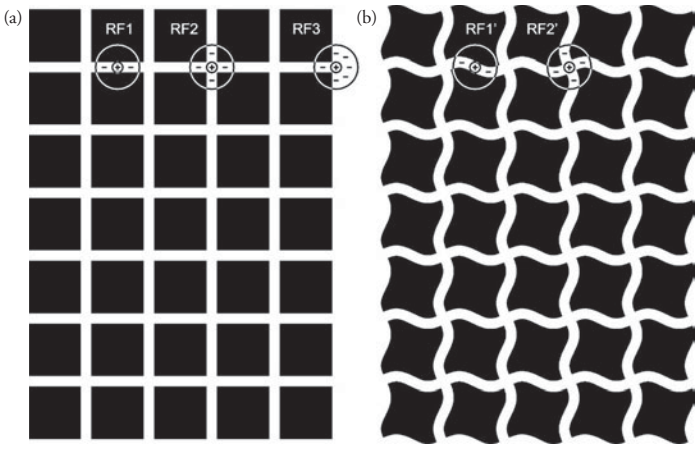


Figure III.45-1. The classical Hermann grid (a) and our Sinusoid version (b). In case of straight streets, illusory spots are seen at the intersections, but they totally disappear in the case of curved streets. The amplitude of the sine curve is less than 10% of its wavelength. RF1, RF2, and RF3 represent the antagonistic receptive fields of retinal ganglion cells. The center of the receptive field is excitatory, and its surround is inhibitory. The classical explanation of the illusion (Baumgartner model) is merely based on the more intensive stimulation of the OFF (inhibitory) surrounds of the receptive fields at the intersections (RF1) compared to the ones inside the street sections (RF2). The inhibitory surrounds projected to the intersections receive white stimulation on four sides, thus eliciting twice as much inhibition as those projected to the middle of the streets, which receive white only on two sides. However, the proportion of stimulation of the OFF surrounds is not influenced by the curvature, since the stimulation of the corresponding receptive fields is equal (RF1 = RF1' and RF2 = RF2'). Therefore, the Baumgartner model would predict spots even in case of the curved grid, which contradicts human perception. We also note that no spots are seen at the end of the streets, where they run into the white background. However, the Baumgartner model would predict even stronger spots there, since RF3 receives even more inhibition than RF2 (not to mention an RF on a homogeneous white paper). (Geier, J., Bernáth, L., Hudák, M., & Séra, L. (2008). Straightness as the main factor of the Hermann grid illusion. *Perception, 37*, 651–665.)

2009). According to this explanation, the center of the receptive field is excitatory, while its surround is inhibitory. Light projected to the center increases while light projected to the surround decreases the activity of the particular ganglion cell. According to the Baumgartner model, the spots at the intersections are caused by the different activity of those ganglion cells, whose receptive fields are located at the intersections (RF1 in Figure III.45-1a) and, of those, whose receptive fields lie in the street sections (RF2). Although the stimulation of the centers of RF1 and RF2 is the same, their surrounds receive different amounts of inhibition (RF1 from four directions, while RF2 from only two directions). As a consequence, the output signal of RF1 will be smaller than that of RF2. This difference is supposed to occur in the perceptual experience as a dark spot at the intersection.

Our refutation of the lateral inhibition–based explanation can be seen in Figure III.45-1. If the cause of the spots were indeed the larger proportion of lateral inhibition at the intersections, then the same illusory spots should be perceived in the curved grid as in the straight grid, since the stimulation of the corresponding receptive fields

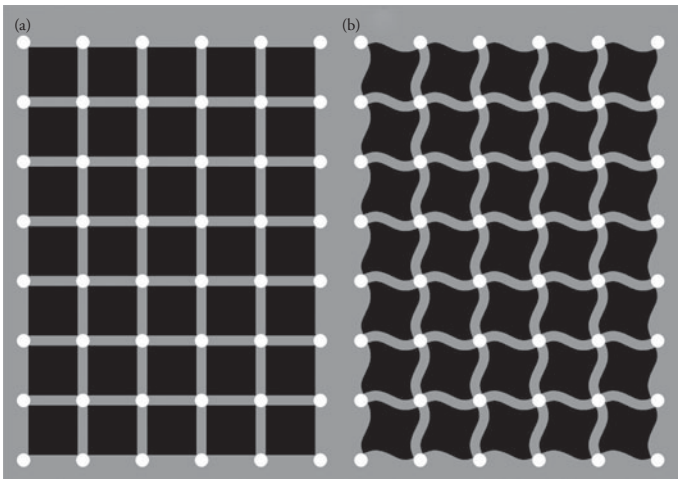
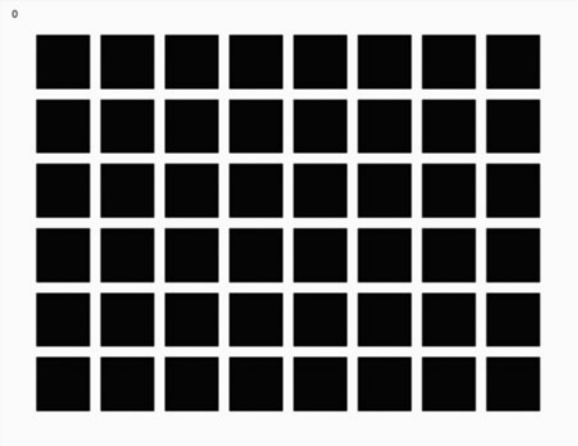
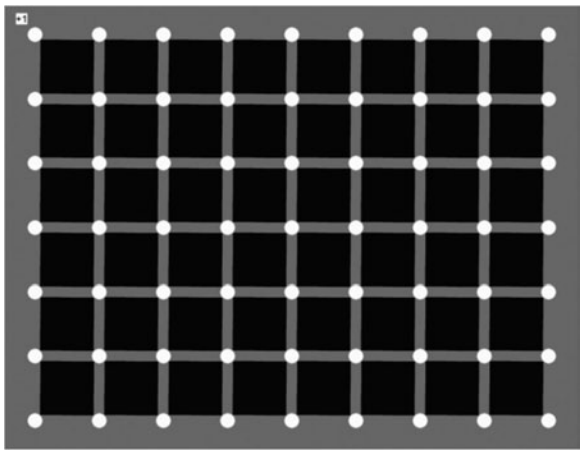


Figure III.45-2. Lingelbach's scintillating grid (a) and our Sinusoid version (b). In the case of straight streets, scintillation is seen at the intersections, but it totally disappears in the case of curved streets. The amplitude of sine curve is less than 10% of its wavelength. To date, there is no generally accepted explanation for the Scintillating grid illusion. It is worth noting that no scintillation is seen at the boundaries of the straight grid pattern. (Geier, J., Bernáth, L., Hudák, M., & Séra, L. (2008). Straightness as the main factor of the Hermann grid illusion. *Perception, 37*, 651–665.)

projected to the straight and the curved grids is equal. However, no spots are seen in the curved grid. Therefore, the Baumgartner model, which is merely based on the different stimulation of receptive fields at the intersections and in the street sections, is not tenable. The fact that curvature significantly changes the Hermann grid illusion calls for the rejection of the classical explanation and development of a new one.



Video III.45-1. Dynamic sinusoid Hermann grid. Open the AVI file (the most suitable program for viewing this video is QuickTime Player). It is not necessary to play the video; instead, frames should be viewed one by one. The first frame is the original, undistorted Hermann grid. When proceeding forward, frame by frame, the grid is gradually curved. The amount of curvature is indicated by a number in the top left corner of the video. Proceed further until the illusory spots disappear. At this point, look at the amount of curvature, which indicates your distortion tolerance for this image. This movie was first shown at ECVP 2004 Budapest (Geier et al., 2004).



Video III.45-2. Dynamic sinusoid scintillating grid. Open the AVI file (the most suitable program for viewing this video is QuickTime Player). It is not necessary to play the video; instead, frames should be viewed one by one. The first frame is the original, undistorted scintillating grid. When proceeding forward, frame by frame, the grid is gradually curved. The amount of curvature is indicated by a number in the top left corner of the video. Proceed further until the illusory scintillation disappears. At this point, look at the amount of curvature, which indicates your distortion tolerance for this image. This movie was first shown at ECVP 2004 Budapest (Geier et al., 2004).

Another contradiction with human perception is that no spots are seen at the end of the streets where they run into the white background, although the receptive fields located there (RF3 in Figure III.45-1) receive even more inhibition than those at the intersections (RF1). Therefore the Baumgartner model would predict even darker spots at the boundaries of the grid pattern than at the intersections. Moreover, in case of a homogeneous white stimulation, such as outside of the grid, the inhibitory ring of the receptive field should receive even more stimulation than at the street intersections (RF2) or at the edge of the grid (RF3). Therefore such areas should be perceived as even darker that the intersections, but this is not the case. These contradictions, irrespective of the curvature, can be observed even in the classical Hermann grid, and, in fact, this observation motivated JG's systematic search for such grid variants that refuted the Baumgartner model even within the grid. The sinusoid grid and our other variants are the result of this systematic search.

Since to date there is no generally accepted explanation for the scintillating grid, we have nothing to challenge concerning that illusion. The fact that curvature eliminates the scintillating effect, too, makes it necessary to search for a common background for these illusions and to develop a general theory for these effects. Another observation that suggests a common background behind the Herman grid and the scintillating grid illusions is that neither spots (Hermann grid) nor scintillation (scintillating grid) is seen at the end of the streets. It is the task of the future to investigate this possible common background.

VARIANTS AND RELEVANT PARAMETERS FOR THE EFFECT

Here the question emerges what the "right" explanation could be. In order to develop a new theory, we should become familiar with the conditions of the appearance and disappearance of the illusion. Our psychophysical experiment was planned to reveal the effects of distortions on the illusion, which might provide a viable basis of a new model. The experiment described here was a tool to provide an answer to the basic question: What is the explanation for the illusory spots in the Hermann grid?

To investigate the necessary conditions for the spots to disappear, we applied further distortion types in addition to the sine curves. These variants are illustrated in Figures III.45-3 and III.45-4.

The wavy grid was introduced for testing the role of the colinearity of neighboring intersections, in comparison

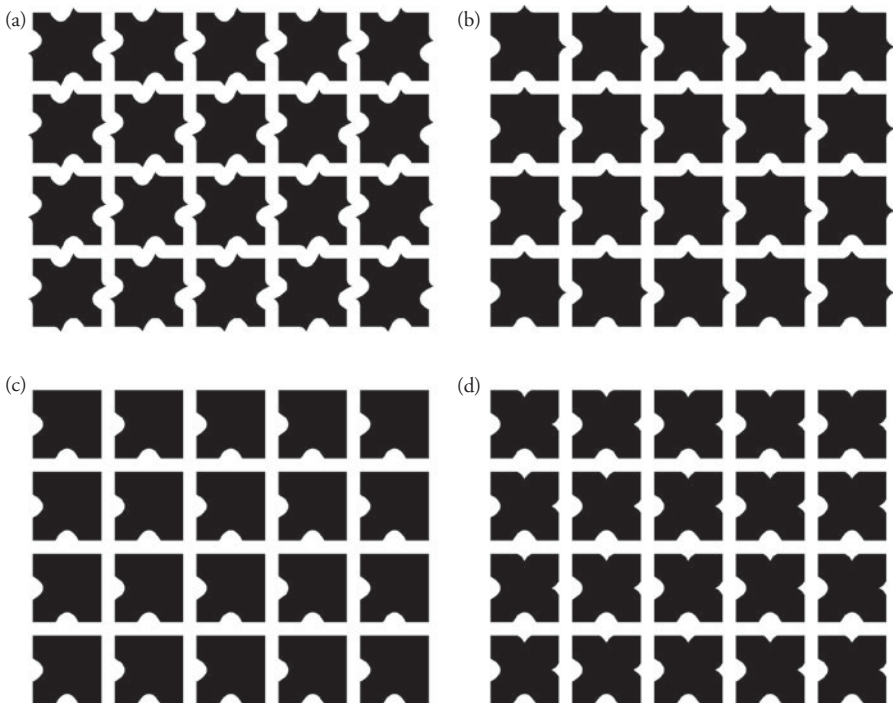


Figure III.45-3. The illustration of four distorted grids: (a) wavy, (b) humped, (c) half-sided humped, and (d) asymmetrical humped. (In the experiment, larger images were used, which consisted of more streets.) The wavy grid comprises of a sine curve in the middle of each street segment, but the neighboring intersections remain co-linear, as opposed to the Sinusoid grid (shown in Figure III.45-1b). The half-sided humped grid was generated from the humped grid by replacing the lower curve of the latter by a straight edge segment. The asymmetrical humped grid was also derived from the humped grid, by flipping its lower curve vertically along the edge of the grid line, so that the lower curves of the street edges were identical. (Geier, J., Bernáth, L., Hudák, M., & Séra, L. (2008). Straightness as the main factor of the Hermann grid illusion. *Perception, 37*, 651–665.)

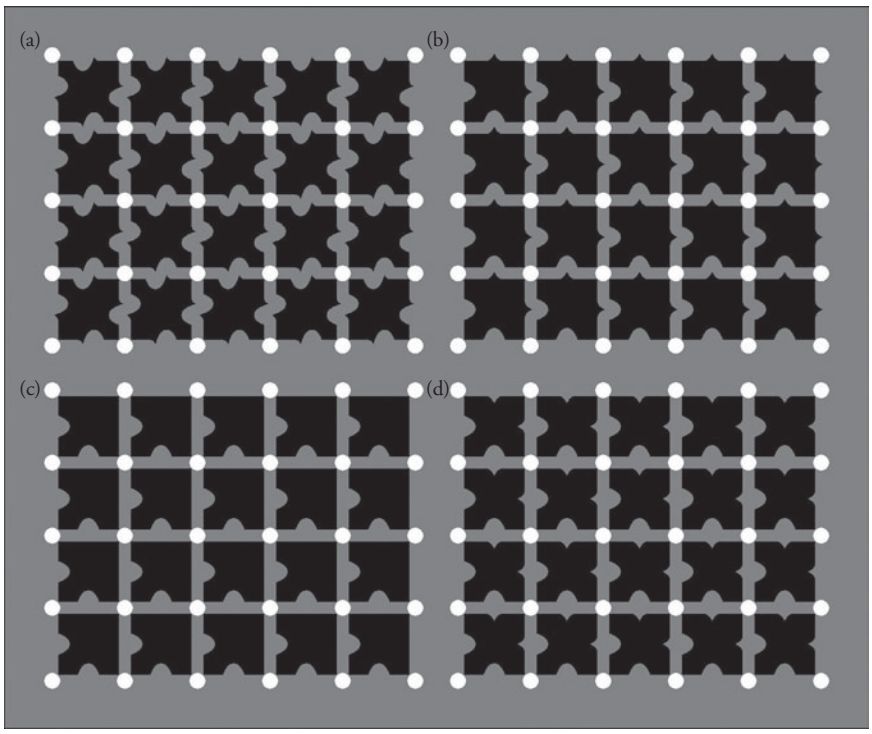


Figure III.45-4. The illustration of four variants of the Scintillating grid. (In the experiment, larger images were used, which consisted of more streets.) This set of images was first shown at ECVP 2004 Budapest (Geier et al., 2004).

with the sinusoid grid (Figure III.45-1b). In the sinusoid grid, not only the straightness of the edges was eliminated by the curvature but also the colinearity of neighboring intersections. The wavy grid, however, preserved the colinearity, while a sine wave was included in the middle part of its street sections. The humped and the half-sided humped grids were introduced to verify whether the straightness of the street edges plays a significant role. The asymmetrical humped and the humped grids tested the effect of black–white area proportions, since these variants differ only in this respect. Thus it was also investigated whether a slight chance has remained for the receptive field explanation.

Earlier experiments concerning the Hermann grid illusion were based on asking the subjects how intensively they perceived the illusory spots (e.g., Wolfe 1984). However, since we can eliminate the illusion completely, it seems more plausible to pose a much more objective question: When does the illusion disappear? Determining the limit when the spots are no longer visible is an even more reliable method than rating the subjective intensity on a scale.

Our psychophysical experiments were based on a measure that we introduced as distortion tolerance (Geier, Bernáth, Hudák, & Séra, 2008). We define the term "distortion tolerance" as the amplitude of curvature at which the illusory spots disappear. Our aim was to reveal, by means of empirical data, on what parameter distortion tolerance depended.

Our experimental results (described in Geier et al., 2008) demonstrate that the distortion tolerance of the distortion type half-sided hump is significantly higher than that of the others. This means that a higher amplitude of curvature is necessary to eliminate the illusion when one side of the streets remains straight. Since it is only the half-sided humped grid that contains straight-edged street sections, it can be concluded that the main cause of the Hermann grid illusion is the straightness of the black–white edges of the streets.

Colinearity of neighboring intersections and black–white area proportions seem unimportant with regard to the illusion. No difference was obtained between the sinusoid grid and the wavy grid, which tested for the role of colinearity. Neither did the humped or the asymmetrical-humped grids differ from each other, by means of which black–white area proportions were controlled. We also tested different street widths, which had no effect.

We conducted the same psychophysical experiments with the scintillating grid. The results were similar: the distortion tolerance was significantly higher in case of the half-sided hump, and the other factors had no effect. Comparing the distortion tolerance of the Hermann grid and the scintillating grid, we obtained approximately twice as large a means in the latter.

This result, that the straightness of the streets is the main factor of these illusions, is likely to be a basis of a new theory, which accounts for both the presence (straight grids) and the absence (curved grids) of the spots and scintillation in a common framework.

DISCUSSION OF RESEARCH AND THEORIES

As expected, an essential principle emerged from the experiment. It seems that the straightness of the streets' edges lies behind the perception of illusory spots and scintillation. The basic idea of our theory for the Hermann grid illusion presented in Geier et al. (2008) stems from these findings.

Having analyzed the distortion types and their effects, we can make an essential but not trivial remark: the location of the change in the illusion is far from the location of the physical change.

In all distortion types, the curving of streets is applied exclusively to the middle part of each street section. Nonetheless, the appearance and disappearance of the

illusory spots and the scintillation take place at the intersections, where no physical change is applied. How does the change in the middle part of the street sections get to the intersections then?

Our answer is that, in fact, it does not even get there. As for the Hermann grid, we suggest that instead of the spots being darker, it is the street segments that are perceived as lighter than the intersections. Our results show that local explanatory principles, such as the antagonism of the receptive fields, are insufficient to provide a common explanation for these phenomena. Instead, long-range interactions between edges and farther areas should be supposed. Next we suggest a possible mechanism for such a long-range interaction.

As a unified explanation of the presence of spots in the classical Hermann grid, and their absence in the curved grid, we propose that a diffusion-like activation-spreading mechanism should be sought behind these phenomena (see Geier et al., 2008, for details).

The small black–white edge segments (the "points" that constitute the edges of the streets) are modeled as small antennas, which radiate inhibitory (black) signals toward the direction of their black side and excitatory (white) signals toward the direction of their white side. These excitatory and inhibitory signals spread along in the close-woven tissue (e.g., the retina) perpendicularly to the edge segments with a small deviation, and therefore they fill in the areas of the image. Since in the classical Hermann grid no edge segment is perpendicular to the intersections themselves, no excitatory signal will reach the intersections. On the other hand, the street sections will receive many excitatory signals, since all edge segments radiate perpendicularly to the street sections. Therefore the street sections will be perceived as much lighter than the intersections. In the curved grids, however, this difference in the amount of excitation received by the streets and the intersections is negligible, since the radiation of many edge segments will be able to reach to the intersections as well as to the street segments, due to their changed angle. Therefore, not only will the streets be "whitened" by the excitatory radiation (which is perpendicular to each edge segment, now positioned in many different angles) but the intersections as well, which causes the dark spots to disappear.

Although we suggested a unified qualitative theory for the classical Hermann grid and its curved versions, the scintillating grid and its curved variants need further explanatory principles. Yet not even the original scintillating grid illusion has been explained in the literature. Nonetheless, we can conclude that the mechanism behind the Hermann grid and the scintillating grid illusions must closely be related. We base this suggestion on our findings that the same distortions make the scintillation disappear that cease the Hermann spots. Moreover, straightness plays an essential role in both phenomena, since both illusions are difficult to eliminate if one side of the streets remains straight. Finally, no scintillation occurs at the end of the streets, similarly to the lack of Hermann spots there.

In conclusion, it seems that the sufficient and the necessary conditions of the Hermann grid illusion and that of the scintillating grid illusion largely overlap. We proposed a unified explanatory principle for the presence and absence of the Hermann grid illusion. The task to extend it to the scintillating grid phenomena, or to provide a better unified theory for all these illusions and non-illusions, remains that of future researchers.

ACKNOWLEDGMENTS

The contribution of MH to this chapter was supported by Grant TÁMOP–4.2.2.B-10/1--2010-0009.

REFERENCES

Baumgartner, G. (1960). Indirekte Größenbestimmung der rezeptiven Felder der Retina beim Menschen mittels der Hermannschen Gittertauschung [Abstract]. *Pflugers Archiv für die gesamte Psychologie, 272*, 21–22.

Bergen, J. R. (1985). Hermann's grid: New and improved [Abstract]. *Investigative Ophthalmology & Visual Science, Suppl., 26*, 280.

Blake, R., & Sekuler, R. (2006). *Perception*. Boston: McGraw-Hill.

Geier, J. (2001). Tranziensek a retinán: A Hermann-grid-illúzió szokásos magyarázatának kritikája. Nyolcadik Magyar Látás Szimpózium, Szeged (2001, szeptember 8, http://www.pszichologia.hu/hirdetes/latabs.html#human). (English translation: Transients in the retina: The critique of the usual explanation of the Hermann grid illusion. http://www.geier.hu/Hermann/firstenglish.html)

Geier, J., Bernáth, L., Hudák, M., & Séra, L. (2008). Straightness as the main factor of the Hermann grid illusion. *Perception, 37*, 651–665.

Geier, J., Séra, L., & Bernáth L. (2004). Stopping the Hermann grid illusion by simple sine distortion. *Perception, 33*(ECVP Abstract Suppl.), 53.

Goldstein, E. B. (2009). *Sensation and perception.* Belmont, CA: Wadsworth Cengage Learning.

Hermann, L. (1870). Eine Erscheinung des simultanen Contrastes. *Archiv für die gesammte Physiologie des Menschen und der Thiere, 3*, 13–15.

Kuffler, S. W. (1953). Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology, 16*, 57–68.

Schrauf, M., Lingelbach, B., & Wist, E. R. (1997). The scintillating grid illusion. *Vision Research, 37*, 1033–1038.

Spillmann, L. (1994). The Hermann grid illusion: A tool for studying human perceptive field organization. *Perception, 23*, 691–708.

Wade, N. J. (2005). Sound and sight: Acoustic figures and visual phenomena. *Perception, 34*(10), 1275–1290.

Wolfe, J. (1984) Global factors in the Hermann grid illusion. *Perception, 13*, 33–40.

Chapter 46
The Staircase Gelb Illusion

Alan Gilchrist

Research on the perception of surface whites, blacks, and grays (i.e., lightness perception) has resulted in many delightful illusions. One of the most dramatic of these is the staircase Gelb illusion (Cataliotti & Gilchrist, 1995), inspired by Gelb's (1929) demonstration that a piece of black paper appears white when suspended in midair and illuminated by a spotlight but black when it is surrounded by white paper also in the spotlight. Figure III.46-1 represents an attempt to capture the illusion in a photograph.

A square piece of matte black paper is suspended in midair and illuminated by a spotlight or a slide projector. If the intensity of illumination reaching the square from the spotlight is about 30 times that of the general room illumination itself, the square, though physically black, will look white. This could hardly be otherwise, given that the black square in the spotlight reflects exactly the same amount of light and makes exactly the same retinal image as a nearby white square not in the spotlight.

Next a second square (actually dark gray) is presented adjacent to the (actually) black square and sharing the same spotlight. This new square now appears completely white, and it makes the first square darken to light gray. Then a third square (actually middle gray) is added next to the second square. This square now appears completely white while both the first and second squares become darker gray. A fourth (light gray) square is added, and finally a white square is added, all within the beam of the spotlight. As each new brighter square is added, it appears white and causes the prior squares to darken.

There is a further illusion. When the final white square is added to the row of squares, the original black square appears as a light middle gray. Thus, although the row of five squares actually spans the entire scale from white to black, the perceived gray shades span a range less than half of that, from white to almost, but not quite, middle gray. This can be seen in Figure III.46-2, where perceived reflectance (lightness) is plotted against actual reflectance. This illusion demonstrates two important aspects of lightness perception: (a) the highest luminance rule of anchoring and (b) gamut compression.

HIGHEST LUMINANCE RULE

The intensity of light reaching one's eye from a gray surface depends on both the percentage of light reflected by the surface (called reflectance) and the intensity of illumination falling on the surface. Given this confound, any shade of gray can reflect any amount of light (called luminance). Thus the luminance reaching one's eye from a surface, by itself, provides absolutely no information about the reflectance of the surface. For this reason it has been generally assumed that perceived lightness depends on relative luminance (Wallach, 1948). However, relative luminance values can produce only relative lightness values (Gilchrist et al., 1999). Yet our percepts are not merely of relative lightness. We perceive specific lightness values. Thus the computation of specific lightness values (say in Munsell units or in perceived reflectance) requires, in addition to relative luminance, an anchoring rule—a rule by which some value of relative luminance is anchored to the black–gray–white scale.

While some have suggested that the average luminance within a scene is anchored at middle gray (Buchsbaum, 1980; Helson, 1964), the empirical evidence suggests that the highest luminance within a scene is anchored at white (Land & McCann, 1971; Li & Gilchrist, 1999). This explains why each new square changes all the lightness values. Each new square increases the highest luminance, making each of the other squares appear darker because the luminance of each square is now farther from the anchor.

Changing the lowest luminance in the spotlight does not have such an effect. That is, if we start with the brightest square (which is white) and successively add darker and darker squares, the addition of the darker squares does not change the perceived lightness of the prior squares.

GAMUT COMPRESSION

The fact that the range of perceived gray shades is so much less than the actual range of grays presents a major challenge to theories of lightness perception. But it may also provide a clue to the solution of the classic problem of lightness constancy. Lightness constancy refers to the empirical fact that the perceived lightness of a gray surface remains remarkably stable even when the luminance of the surface changes due to a change in the illumination falling on that surface. Various theories have offered an explanation of lightness constancy. Each theory also contains implications for the predicted pattern of failures of constancy. Though

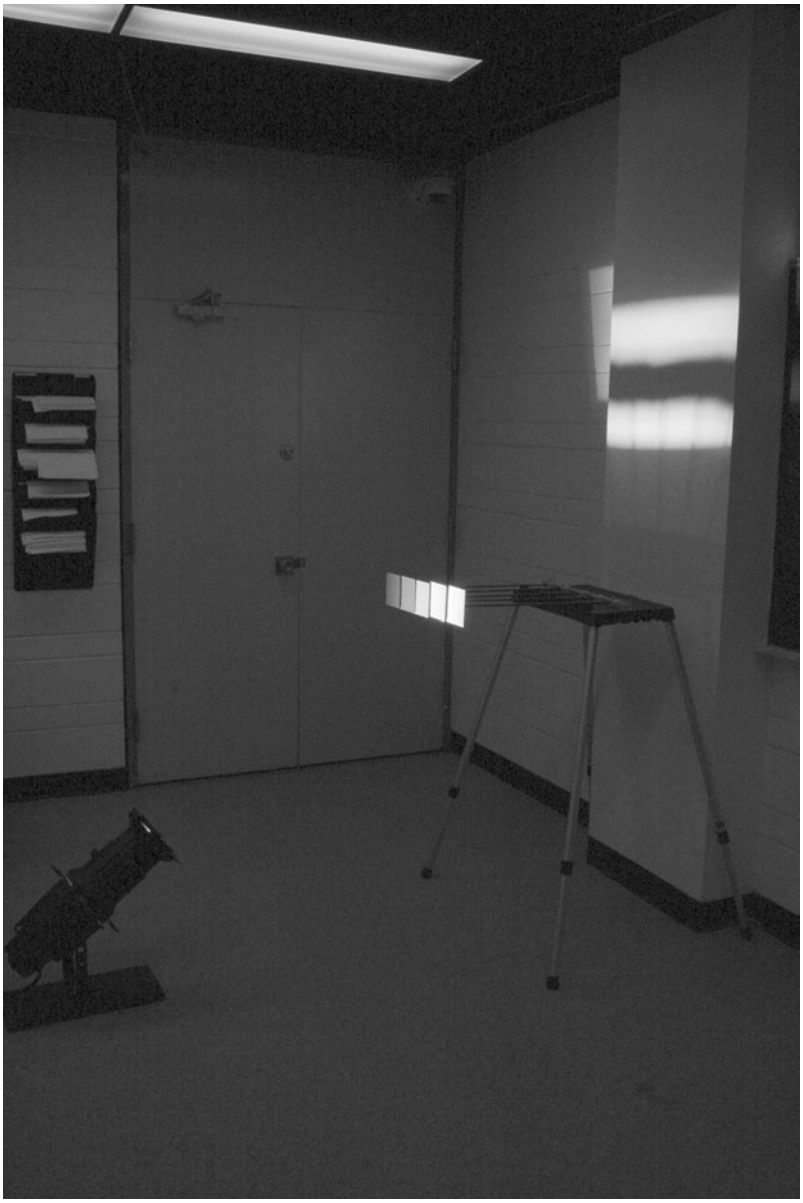


Figure III.46-1. A row of five squares ranging from black to white is suspended in midair and illuminated by a spotlight 30 times brighter than room illumination. The perceived range of grays is dramatically compressed relative to the actual values, the perceived range extending only from light middle gray to white. (Photo: Gene Anderson.)

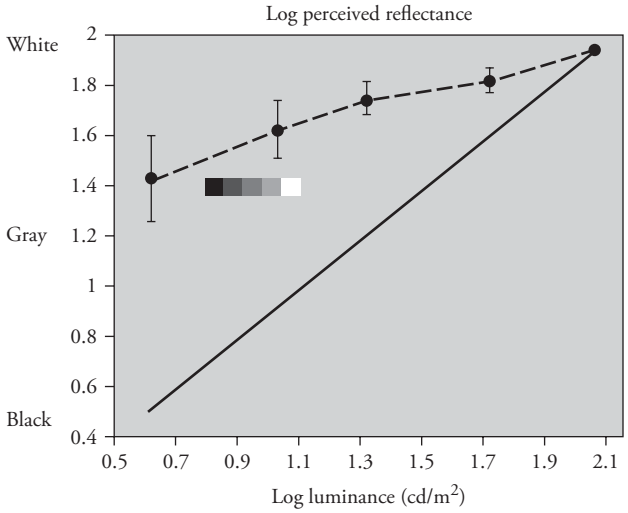


Figure III.46-2. When the perceived lightness values of the five squares are plotted against their actual reflectances, the resulting slope is quite shallow relative to the veridical slope. (Adapted with permission from Gilchrist et al., 1999.)

these implications are typically not explicit, they may offer the best means for evaluating theories.

As with other perceptual constancies, such as size constancy and shape constancy, lightness constancy is never perfect. In addition, the particular way in which constancy fails must be highly diagnostic of the lightness software embodied in the human visual system. To a first approximation, the lightness of a surface will remain constant when moved into a spotlight or a shadow, but in fact it will appear slightly lighter in the spotlight and slightly darker in the shadow. More important, a closer look at the data reveals that not all shades of gray show the same degree of failure. In a region of high illumination (such as a spotlight), dark grays are lightened more than light grays, while in a shadow, the opposite occurs—light grays are darkened more than dark grays. In other words, when a variety of gray surfaces are placed within a spotlight, the range of perceived gray shades is upwardly compressed (relative to the actual range). The same thing happens when the gray shades are placed within a shadow, except that the range is compressed downward rather than upward.

This means that the extreme gamut compression found with the staircase Gelb illusion is not merely an isolated feature of a special case. Rather it seems to be a universal form of lightness constancy failures (Gilchrist, 1980, 2006; Schirillo, Reeves, & Arend, 1990). Some degree of gamut compression must occur whenever a scene contains more than one level of illumination, a feature of almost every scene we observe. Thus the staircase Gelb illusion not only presents a challenge to lightness theory but strongly constrains such a theory as well.

MODELING THE COMPRESSION

Among the most popular theories of lightness are what have been called decomposition models. According to the logic of inverse optics, the retinal image is decomposed into the factors, such as illumination and reflectance, that originally combined to produce the image. Bergström (1977), for example, proposed a vector-analytic method by which the retinal image is parsed into common and relative components of reflected light, decomposing the image into patterns of reflectance, illumination, and three-dimensional form. Adelson and Pentland (1996) offered an elegant workshop metaphor that seems functionally equivalent to Bergström. Although these models are primarily models of veridical perception, they could be extended to failures of lightness constancy merely by postulating a failure in the decomposition. Such a failure would result in a darkening of patches within a shadow, for example. But all shades of gray would be darkened by the same amount. Thus the decomposition failure would not produce gamut compression.

Intrinsic image models (Arend, 1994; Barrow & Tenenbaum, 1978; Gilchrist, Delman, & Jacobsen, 1983) suggest that edges in the retinal image are classified as changes in either reflectance or illumination. Imperfections in such edge classification, as proposed for example by Ross and Pessoa (2000) and Gilchrist (1988), would make surfaces in a shadow appear darker, but, once again, all shades would be darkened by the same amount, with no compression.

The only explicit theory of failures of lightness constancy is found in the concept of codetermination proposed by Kardos (1934). Koffka (1935) emphasized that lightness, like size and shape, is seen relative to a frame of reference (essentially a field of illumination). Kardos embraced the concept of frameworks but argued that there is crosstalk between frameworks (Sedgwick, 2003), proposing that the lightness of a target surface represents a compromise between the lightness of that target relative to the framework to which it belongs and the lightness of that target relative to a surrounding or an adjacent framework. Indeed, the empirical literature shows that perceived lightness always lies between these two values; at least it is hard to think of an exception. Thus codetermination offers a potential explanation for gamut compression.

Several models (Bressan, 2006; Gilchrist, 2006) have attempted to further develop the concept of codetermination. The retinal image is parsed into adjacent frameworks of illumination, and lightness is determined by a weighted average of values computed in a local framework and a foreign, surrounding, or global framework. Gilchrist and Bressan have argued that such a weighted average would produce gamut compression. If the spotlight is bright enough (about 30 times brighter than room illumination) so that a black surface in the spotlight has the same luminance as a white surface outside the spotlight, then the luminance of each of the five squares will be equal to or higher than that of a white square in the surrounding room (global framework). Thus if all five squares are computed to have the same global value (white), and those equal values are averaged together with the stepwise values computed in the local framework, the net result is compression.

However, this scheme implies that increasing the spotlight intensity beyond this ($30x$) value will not increase the compression any further, given that the global values are already maxed out. Yet empirical data show that compression continues to increase beyond this value of spotlight intensity.

RELATIVE ILLUMINATION

One thing is clear, however: the gamut compression depends on the relative illumination between the five squares and the surrounding room. We know that there is no compression when the illumination on the five squares is equal to that of the room. Nor would there be any compression if the five squares were to fill the observer's entire visual field, for example, by painting those shades onto the interior surface of a hemisphere placed around the observer's head. The sheer fact that gamut compression requires a juxtaposition of different spatial fields of illumination resonates with the concept of codetermination, even though the precise formula for codetermination remains unknown.

FACTORS THAT AFFECT COMPRESSION

Could the compression not result merely from the traditional concepts of bright and dark adaptation? The spotlight extends the dynamic luminance range within the room, and perhaps the visual system attempts to accommodate this extended range by compressing its response, especially at the high and low extremes (Radonjic, Allred, Gilchrist, & Brainard, 2011). The five-square display constitutes a sort of island of bright light surrounded by the dimmer illumination of the room. If the eyes are adapted to the general level of room illumination, perhaps the perceptual response to the five squares is compressed due to their outlying luminance values.

However, a simple variation dispels this hypothesis. If a narrow white border is placed around the five-square display, the compression vanishes and the squares are seen almost veridically. Although photographs fail to convey the scene faithfully, Figure III.46-3 gives a good idea of it. Compression does not occur even though the luminance values within the spotlight are still outliers compared to the dimmer room (Gilchrist et al., 1999). Surrounding the display with a black border does not have this effect.

It might be thought that the white border would darken the squares due to local contrast, but this does not seem to be the case. It is not necessary that the white border make direct contact with any of the squares—only that the border totally surrounds them. This can be seen in

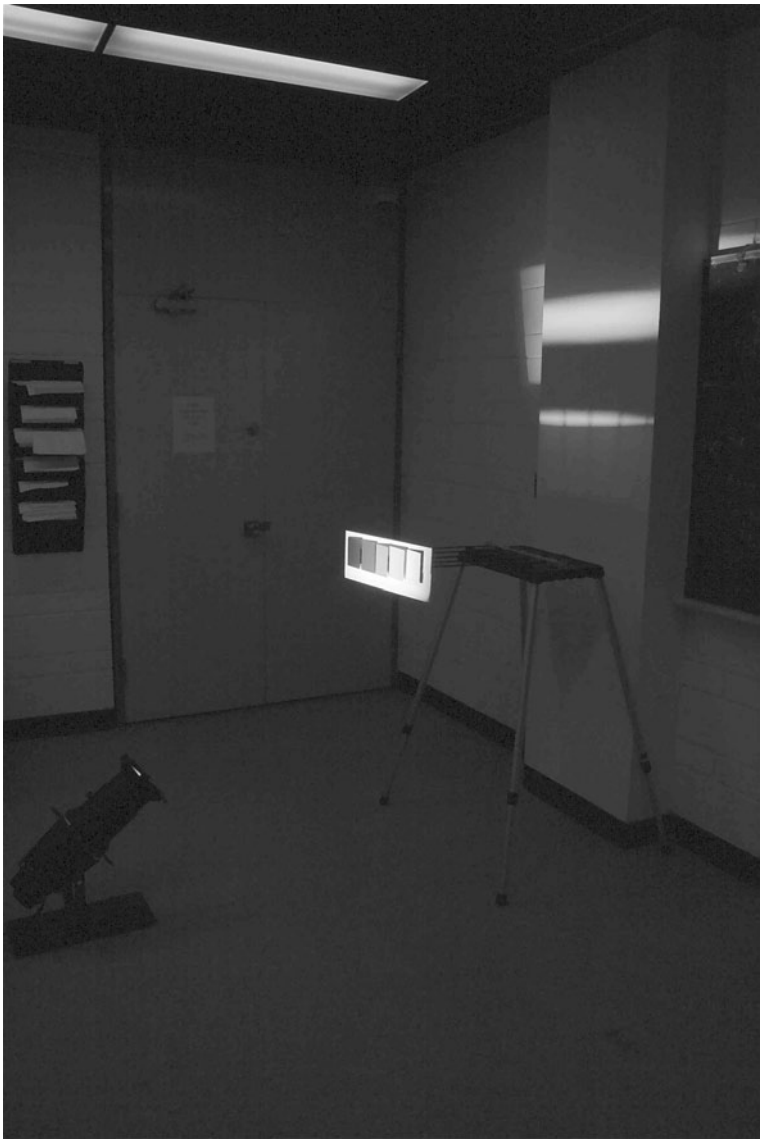


Figure III.46-3. When a white border is placed around the five squares (but within the spotlight), the compression vanishes. The white border seems to insulate the squares from the influence of the surrounding room. This fact contradicts potential explanations of the compression based on overall luminance range, or faulty adaptation. (Photo: Gene Anderson.)

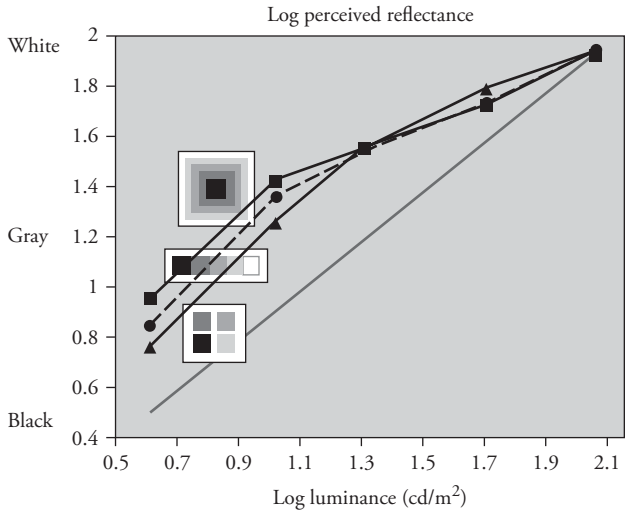


Figure III.46-4. To eliminate the gamut compression by surrounding the squares with a white border, it is not necessary that the border be in contact with each square—only that it surround the squares. This fact suggests that the insulation effect cannot be reduced to local contrast (Adapted with permission from Gilchrist et al., 1999.)

Figure III.46-4, which shows the results from several configurations that have been tested in the spotlight. In the window panes configuration (bottom of three icons), the black square is in direct contact with white along its entire border, while in the bullseye configuration (top), the black square has no direct contact with the remote white border. Although there does seem to be a weak trend consistent with a local contrast effect, these data revealed no significant difference in perceived lightness for any of the squares (including the black square) across conditions.

Somehow, when the configuration in the spotlight is completely surrounded by a white border, the border seems to insulate the patches in the spotlight from the influence of the larger framework of the room. Herein lies an inviting thesis topic.

ARTICULATION

The degree of compression in the staircase Gelb illusion is very sensitive to the number of patches in the spotlight. This factor has been called articulation, and it goes back to the pioneering lightness work of David Katz (1935).

Katz (1935) is widely known for his phenomenological analysis of the modes of color appearance. He is less well known for establishing the basic psychophysical methods for the study of lightness. Among those methods is that of asymmetric matching, in which an observer views two gray disks, one located in an illuminated field and the other in an adjacent shadowed field. The observer chooses the gray level of one disk that makes it appear equal to the other disk in lightness. Katz was the first to study the factors that influence the degree of lightness constancy in such a configuration. The most potent factor he found he called articulation or diversity, which was operationalized in his experiments by the number of different patches within a field of illumination. He found that the degree of constancy was very low when articulation was low but almost perfect with 48 patches in each field.

Because compression of the five shades seen in the spotlight represents a failure of constancy, it is not surprising that increasing the number of patches in the spotlight reduces the compression. This can be seen in Figure III.46-5, which shows observer matches (between subjects) for Mondrian displays with differing levels of articulation (Gilchrist et al., 1999).

Currently there are a number of theories of lightness constancy, including high-level decomposition and inferential theories; midlevel Gestalt, framework theories; and low-level theories based on peripheral sensory mechanisms. Often these theories can equally account for lightness constancy data. But, as I have argued (Gilchrist, 2006), failures of constancy may offer a more rigorous test of theories. Every theory of lightness implies a particular pattern of failures of constancy. No two theories predict the same pattern of failures, unless the theories are functionally equivalent.

Thus the question of how to account for the gamut compression found in the staircase Gelb illusion becomes a compelling one. Its importance is even further enhanced by the fact that gamut compression turns out to be the universal signature of lightness constancy failures. After 150 years of lightness research, this has only recently been recognized, although the reason for such late recognition is simple. Gamut compression will not be revealed by standard asymmetric matching methods or by nulling methods. It can be revealed only when targets of multiple gray levels are matched using a Munsell chart or other similar scale. This has rarely been done.

Thus while this universal form of lightness constancy failures (gamut compression) presents one of the most important challenges to a theory of lightness, it also appears to offer one of the biggest clues. Attempts to model this pattern are likely to be well rewarded.

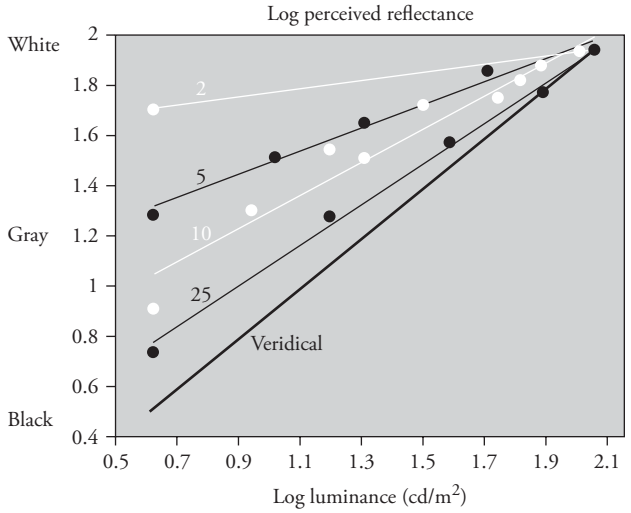


Figure III.46-5. The amount of compression strongly varies with the number of squares in the spotlight, a factor Katz (1935) called articulation. The graphs show regression lines for Mondrian-type groups of squares in the spotlight consisting of 2, 5, 10, and 25 squares, each with the same range from black to white. (Adapted with permission from Gilchrist et al., 1999.)

Acknowledgment: Supported by NSF: BCS-1230793.

REFERENCES

Adelson, E. H., & Pentland, A. P. (1996). The perception of shading and reflectance. In D. Knill & W. Richards (Eds.), *Perception as Bayesian inference* (pp. 409–423). New York: Cambridge University Press.

Arend, L. (1994). Surface colors, illumination, and surface geometry: Intrinsic-image models of human color perception. In A. Gilchrist (Ed.), *Lightness, brightness, and transparency* (pp. 159–213). Hillsdale, NJ: Lawrence Erlbaum.

Barrow, H. G., & Tenenbaum, J. (1978). Recovering intrinsic scene characteristics from images. In A. R. Hanson & E. M. Riseman (Eds.), *Computer vision systems* (pp. 3–26). Orlando, FL: Academic Press.

Bergström, S. S. (1977). Common and relative components of reflected light as information about the illumination, colour, and three-dimensional form of objects. *Scandanavian Journal of Psychology*, *18*, 180–186.

Bressan, P. (2006). The place of white in a world of grays: A double-anchoring theory of lightness perception. *Psychological Review*, *113*, 526–553.

Buchsbaum, G. (1980). A spatial processor model for object color perception. *Journal of the Franklin Institute*, *310*, 1–26.

Cataliotti, J., & Gilchrist, A. L. (1995). Local and global processes in lightness perception. *Perception & Psychophysics*, *57*(2), 125–135.

Gelb, A. (1929). Die "farbenkonstanz" der sehdinge [The color of seen things]. *Handbuch der normalen und pathologischen Physiologie*, *12*, 594–678.

Gilchrist, A. L. (1980). When does perceived lightness depend on perceived spatial arrangement? *Perception & Psychophysics*, *28*(6), 527–538.

Gilchrist, A. (1988). Lightness contrast and failures of constancy: A common explanation. *Perception & Psychophysics*, *43*(5), 415–424.

Gilchrist, A. (2006). *Seeing black and white*. New York: Oxford University Press.

Gilchrist, A., Delman, S., & Jacobsen, A. (1983). The classification and integration of edges as critical to the perception of reflectance and illumination. *Perception & Psychophysics*, *33*(5), 425–436.

Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., … Economou, E. (1999). An anchoring theory of lightness perception. *Psychological Review*, *106*(4), 795–834.

Kardos, L. (1934). Ding und schatten [Object and shadow]. *Zeitschrift fur Psychologie*, *23*, 1–184.

Katz, D. (1935). *The world of colour*. London: Kegan Paul.

Koffka, K. (1935). *Principles of Gestalt Psychology*. New York; Harcourt, Brace, and World.

Land, E. H., & McCann, J. J. (1971). Lightness and retinex theory. *Journal of the Optical Society of America A*, *61*, 1–11.

Li, X., & Gilchrist, A. (1999). Relative area and relative luminance combine to anchor surface lightness values. *Perception & Psychophysics*, *61*(5), 771–785.

Radonjic, A., Allred, S., Gilchrist, A., & Brainard, D. (2011). The dynamic range of human lightness perception. *Current Biology*, *21*(22), 1931–1936.

Ross, W., & Pessoa, L. (2000). Lightness from contrast: A selective integration model. *Perception & Psychophysics*, *62*(6), 1160–1181.

Schirillo, J. A., Reeves, A., & Arend, L. (1990). Perceived lightness, but not brightness, of achromatic surfaces depends on perceived depth information. *Perception& Psychophysics*, *48*(1), 82–90.

Sedgwick, H. (2003). Relating direct and indirect perception of spatial layout. In H. Hecht, R. Schwartz, & M. Atherton (Eds.), *Looking into pictures: An interdisciplinary approach to pictorial space*. Cambridge, MA: MIT Press, pp. 61–75.

Wallach, H. (1948). Brightness constancy and the nature of achromatic colors. *Journal of Experimental Psychology*, *38*, 310–324.

Chapter 47
The Breathing Light Illusion

Illusory Size and Brightness Variation Induced by Motion

Simone Gori, Enrico Giora, and D. Alan Stubbs

This chapter is in memory of D. Alan Stubbs, a wonderful friend and colleague who will be never forget.

INTRODUCTION

The aim of this chapter is to present a size and brightness illusion called the breathing light illusion (BLI; Gori & Stubbs, 2006) and to review studies that help to understand the underlying mechanisms that could be responsible for that phenomenon.

It is often assumed that perceiving a surface as a source of light depends just on its physical radiant emission (see Kingdom, 2011, for a review regarding luminance perception). However, since the pioneering investigations on simultaneous contrast by the Persian natural philosopher Ibn Al-Haytham (c. 965–1040 ad), also known as Alhazen, the subjective nature of color sensation has been highlighted (Wade, 1996). Arguing that color appearance was partly due to a mental process, Alhazen anticipated theories later developed by Helmholtz (1867) and Hering (1878; also see Kingdom, 1997). Interestingly, Renaissance engravers knew empirically how to modulate the brightness by manipulating the edges' appearance (Zavagno & Massironi, 2006). Classical investigations (Heinemann, 1955; Hess & Pretori, 1894; Jameson & Hurvich, 1964; Wallach, 1948) confirmed that brightness is affected by contextual properties' modulation of a pattern.

Although models based on simple, low-level mechanisms were normally used to explain effects such as border contrast or Mach bands (Mach, 1865), more complex mechanisms occurring later in the visual system must be taken into account when one considers brightness effects influenced by the perceptual organization of the stimuli (Adelson, 1993; Agostini & Galmonte, 2002b; Agostini & Proffitt, 1993; Benary, 1924; Bressan, 2006; Economou, Zdravković, & Gilchrist, 2007; Gilchrist et al., 1999; Kingdom, 2008; Todorović, 2006; Troncoso, Macknik, & Martinez-Conde, 2005; Vasarely, 1970; Zdravković, Economou, & Gilchrist, 2012). Phenomena such as illusory contours, area contrast, color constancy, depth planes, coherent motion, and texture contrast all show Gestaltic properties and indicate the existence of neurophysiological mechanisms within the visual pathways with long-range interactions (Gori & Spillmann, 2010; Spillmann, 1997, 2009; Spillmann & Werner, 1996).

A peculiar characteristic that influences the final luminance percept concerns the figure boundaries (Cohen & Grossberg, 1984; Kingdom & Moulden, 1988). Two dramatic examples of how the boundaries affect the brightness are the Cornsweet illusion (Cornsweet, 1970) and the watercolor illusion (Pinna, Brelstaff, & Spillmann, 2001). Importantly, a blurred boundary leads to an increment of the brightness independently of the physical luminance value (Gori & Stubbs, 2006; Kennedy, 1976; Spillmann, Hardy, Delahunt, Pinna, & Werner, 2010; Todorović, 2006; Zavagno, 1999).

DESCRIPTION OF THE PHENOMENON

The phenomenon reported in this chapter regards stimuli with blurred boundaries but presented with dynamic viewing (Gori & Stubbs, 2006). The effect (the BLI) arises in a circular white spot on a black background. The boundaries of the spot are characterized by a gradient in luminance that gives an impression of blur (Fig. III.47-1a). The blurred white spot seems to glow and exhibits a self-luminance appearance. A variation, called the dynamic luminance-gradient illusion, is composed of black radial spokes on a white background that become gradually narrower and whiter from the borders to the center of the figure (Fig. III.47-1b). The blurred spot of these two images appears wider, brighter, and more diffuse when it is approached, whereas it seems smaller, darker, and sharper when one recedes from it (Gori & Stubbs, 2006). When one approaches the negatives of Figures III.47-1a and III.47-1b, the illusory effect for size remains the same, whereas the brightness effect reverses.

Other well-known patterns that have similar luminance gradient changes can produce similar phenomena. For example, the Vasarely pattern (Troncoso et al., 2005; Vasarely, 1970) or the glare-effect figure (Zavagno, 1999) elicit the illusory effects described here if one moves toward and away from them. Although both phenomena (i.e., the variation in size and in brightness) can be explained by a

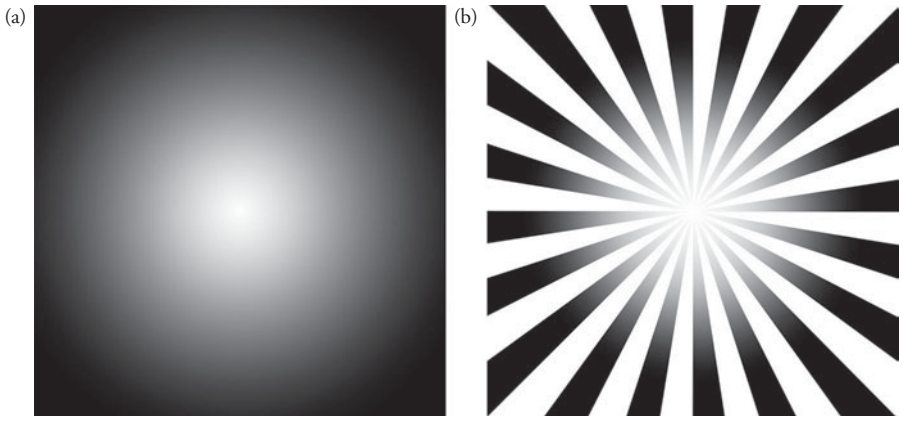


Figure III.47-1. Panel (a): A simple version of the BLI. Panel (b): A gray-scale version of the dynamic luminance-gradient illusion. Approaching these patterns by moving one's head toward them makes the spot appear to become larger, more diffuse, and filled with white. On receding from it, the spot's center remains white but the remainder appears smaller, darker, and sharper. Figures created by Simone Gori and D. Alan Stubbs.

common mechanism, described later, they differ enough to deserve separate coverage.

The variation in size is the first illusory effect elicited by the BLI. Size constancy is a well-known phenomenon (see Ross & Plug, 1998, for a review): when the observer moves toward a stimulus, its retinal image becomes bigger, but the brain normally corrects that and no size variation is experienced. Apparently, the BLI is a quite unique pattern that breaks the size constancy: the blurred spot grows, indeed, with the motion toward the stimulus.

The variation in brightness as a function of the observer's motion is the other illusory effect elicited by the BLI. It resembles the perception that can be experienced moving toward a real light source. The variation in perceived luminance also produces an effect on figures with sharp boundaries superimposed on the center of the BLI (Gori & Stubbs, 2006). In particular, if a small disk is placed on the spot's center and the observer moves backward and forward, the disk will show a variation in brightness. The increased perceived contrast in the superimposed disk seems causally linked with the illusory brightness enhancement (Gori, Giora, & Agostini, 2008). Based on this observation, Gori, Giora, and Agostini (2010) investigated the increment of the achromatic contrast of a gray sharp-boundary disk when it was superimposed on the BLI. This augmentation of the perceived contrast in the dynamic presentation of the BLI was significantly stronger than the effect that Agostini and Galmonte (2002a) obtained in static presentation. Moreover, Gori et al. provided an indirect measurement of the BLI brightness variation.

DISCUSSION

The self-luminance appearance of a blurred spot during a static view was investigated before the discovery of the BLI. Zavagno (1999) described the glare impression produced by a luminance gradient as a bright light source. This brightness increment seems to involve a cortical process and cannot be explained by retinal or lateral geniculate nucleus mechanisms, as shown by Lu, Zavagno, and Liu (2006). These authors reported, indeed, that the duration of a glare-pattern afterimage did not differ from a nonblurred figure afterimage. In a functional magnetic resonance imaging study, Leonards, Troscianko, Lazeyras, and Ibanez (2005) suggested that blurred stimuli producing glare are processed by a region of the extrastriate cortex, specifically in the occipitotemporal and/or collateral sulcus at the border between Brodmann areas 19 and 37.

However, the idea that color perception and figures' surface appearance are critically affected by edge properties has a long history behind (e.g., Kanizsa, 1954, 1957; MacLeod, 1947). Burgh (1964) showed an increase of achromatic contrast in a classical simultaneous-contrast display by blurring the edge of a gray test patch, defocusing the projector at the edge of the target region. The blurred boundaries can affect the brightness of a superimposed figure with sharp boundaries (Agostini & Galmonte, 2002a; Gori et al., 2010; Gori & Stubbs, 2006). Agostini and Galmonte showed that a middle-gray region placed at the center of an area filled by a linear achromatic gradient from black (outer part) to white (inner part) was perceived to be much darker than an identical middle-gray region surrounded by a reversed gradient. This phenomenon appears significantly stronger than the classical achromatic simultaneous-contrast effect. According to the authors, these results cannot be interpreted within the framework of lateral inhibition theories, which claim that the brightness of a region is the result of an inhibition process determined only by the intensity of its surround (Agostini & Proffitt, 1993). Also, the Gilchrist et al. (1999) theory cannot explain this result satisfactorily. A possible explanation is that the luminance gradient in the display in Agostini and Galmonte (2002a) is perceived by the visual system as a change in the illumination level, rather than in the reflectance of the surfaces (see Kingdom, 2008, for a review). This interpretation is congruent with the logic of the *albedo* hypothesis (Kozaki & Noguchi, 1976), which states that there is an invariant relationship between perceived surface brightness and perceived illumination level. An error in perceived brightness accompanies an equal but opposite error in perceived illumination.

In order to find a possible explanation for the new illusory phenomenon represented by the BLI seen in dynamical view, several variables were controlled by Gori and Stubbs (2006). Neither shape, colors, size, nor number of spikes are critical for the illusion in the dynamic luminance-gradient variation. The illusory effects are, indeed, still present no matter how these variables are manipulated (see Figures III.47-2 and III.47-3 for examples).

The one variable that seems critical in the stimulus to produce the BLI is the luminance gradient that leads to

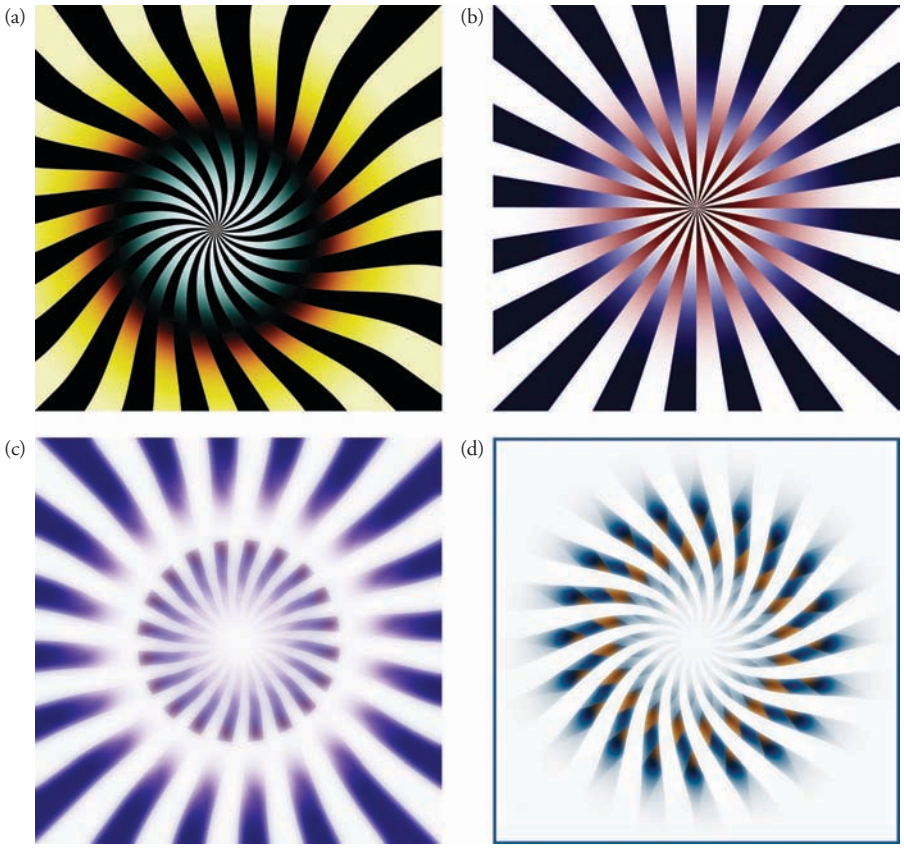


Figure III.47-2. Some variations of the BLI. Approaching the stimuli in panels a, c, and d with central fixation, an illusory rotation is perceived together with the typical illusory effect of the BLI. This illusory rotation is the same experienced in the rotating tilted lines illusion (Gori & Hamburger, 2006; Gori & Yazdanbakhsh, 2008; Yazdanbakhsh & Gori, 2008). Figures created by Simone Gori and D. Alan Stubbs.

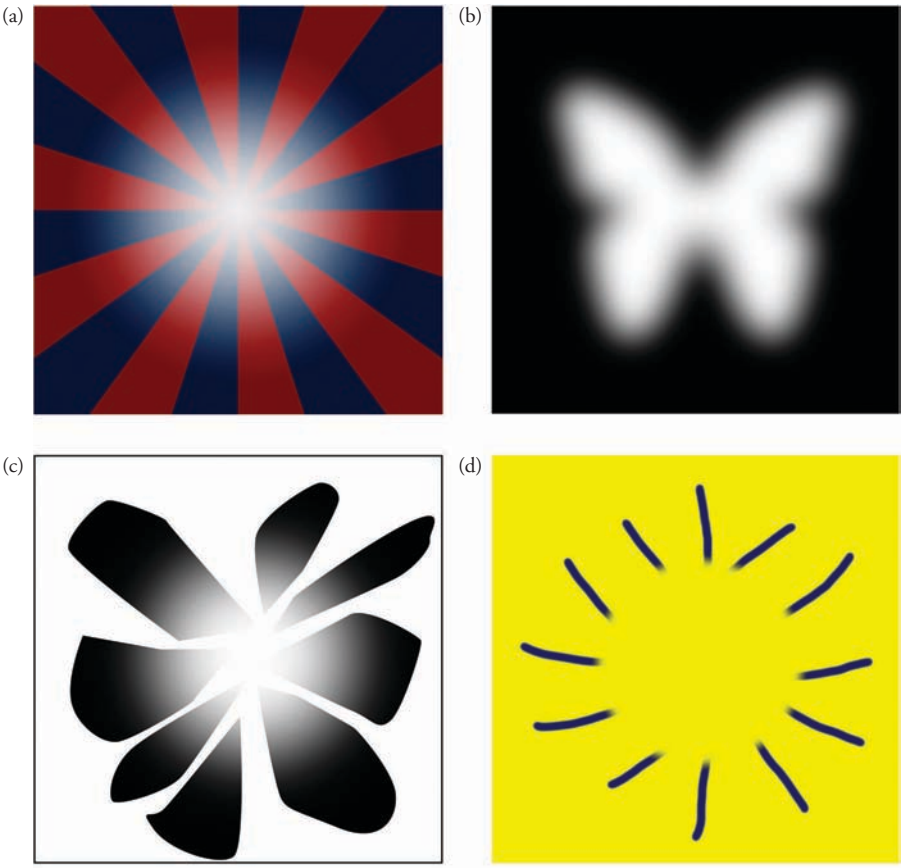


Figure III.47-3. Some variations of the BLI demonstrating that the illusion also occurs with various colors and other shapes (panels b and c); it can also occur with amodal completion (panel c) and when the gradients are minimal (panel d). Figures created by Simone Gori and D. Alan Stubbs.

the appearance of blur. The illusory effects disappear if the boundaries are sharp or luminance changes occur in steps, without any graded change of luminance (Gori & Stubbs, 2006). The crucial role of the luminance gradient was tested by Gori et al. (2010). The augmentation of the perceived contrast on the superimposed disk during dynamic viewing was present when the background was characterized by a luminance gradient (their Experiment 1). However, the enhancement in brightness disappeared when no luminance gradient was shown in the background (their Experiment 2). Assuming the luminance gradient is the crucial variable to obtain both illusory phenomena, the open question remained regarding how this characteristic could cause these effects during dynamic viewing.

Anstis, Gori, and Wehrhahn (2007) proposed an explanation of the illusion in terms of a combination of negative afterimages and the physical stimulus. A negative afterimage is an image that continues to appear in one's vision after the end of the original image exposure. The negative afterimage is principally caused by adaptation of the eye's photoreceptors, which could be very fast and cause several visual illusions (see Petrov & Popple, 2002). During the dynamical observation of the BLI, the negative afterimage of the stimulus created on the retina in the first observer's position will be superimposed on the physical image at the second observer's position. The combination of the afterimage and the physical stimulus would lead to both illusory effects characterizing the BLI. Anstis et al. (2007) presented several examples to prove the crucial involvement of the negative afterimage in explaining the illusion: both increments (size and brightness) disappeared with the suppression of the negative afterimage when one approached the BLI. To directly experience this potential explanation, see Figure III.47-4. A few seconds' fixation with the head at a fixed distance from the pattern in Figure III.47-4a adapts the observer to the luminance of the initial stimulus. Moving the head sharply forward while fixating the cross generates a curious effect: two small illusory blurry disks are briefly perceived. A small, blurry white dot appears on the inner edge of the black dot at 3 o'clock, while a small, blurry black dot appears on the inner edge of the white dot at 9 o'clock. An additional pause serves to readapt the observer to the close-up view of the stimulus. Then, moving the head backward produces a large white flare on the outer edge of the black dot at 9 o'clock and a large black flare on the outer edge of the white dot at 3 o'clock. Figures III.47-4b and III.47-4c show an artist's impression of how these afterimages look. They are clearly generated by the stimulus in which changes in luminance are strongest. Changes are smaller than the stimulus when the head moves forward and larger when the head moves back.

Figure III.47-4, bottom, right, explains these results for a fixated white disk with a Gaussian luminance profile. Figure III.47-4d shows a narrow spatial luminance profile at time T1 when the head is far away and a broader luminance profile at time T2 when the head is closer. In Figure III.47-4e the afterimage from T1 is superimposed on the stimulus at T2 and vice versa. These afterimage profiles are drawn with dashed lines. The resulting perceived profiles, drawn with thick lines, are differences of Gaussians. The upper profile, when the head is far away at T1, has the typical Mexican-hat profile, with the spot apparently sharpened up, and in the lower profile, when the head is in the near position at time T2, the central bright area enlarges and the spot appears diffuse and blurry. This matches exactly the observed percepts when the BLI is dynamically viewed. Thus the combination of the stimulus with its own differently sized afterimage leads to increased spatial differentiation when the observer's head is far away and increased spatial summation when the observer's head is close in.

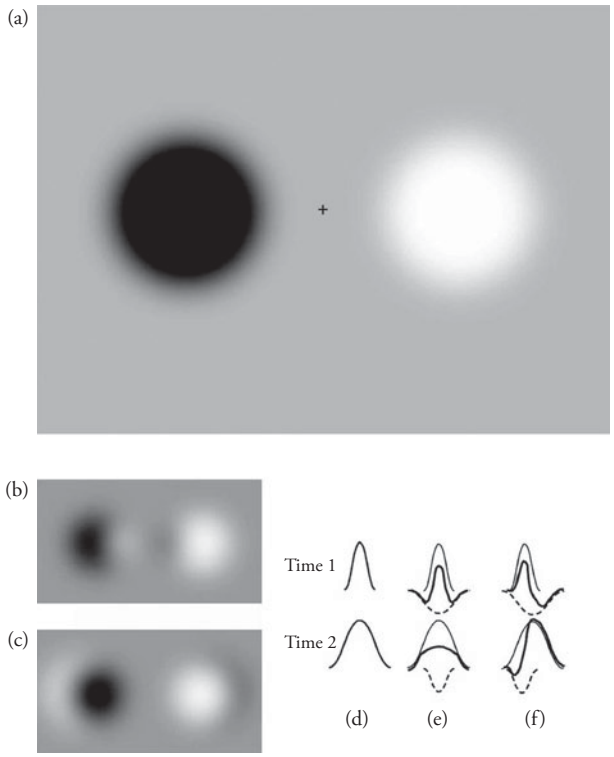


Figure III.47-4. In order to understand the mechanism that likely produces the BLI, try to fixate the cross in panel a for 5 to 10 s. When the head is moved sharply forward to half the viewing distance, the stimulus briefly looks like in panel b, with small, blurry afterimages superimposed. After few seconds of fixation, if the head is moved back again, the stimulus will briefly look like in panel (c), with large, blurry afterimages. Panels d, e, and f represent the hypothetical retinal images of blurry patterns at different times and head positions. Figures created by Stuart Anstis, Simone Gori and Christian Wehrhahn.

With eccentric fixation upon the cross and the white disk to the right of fixation, the afterimage from T2 will be shifted laterally upon the stimulus at T1 and vice versa. Figure III.47-4f shows the displaced afterimage profiles drawn with dashed lines and the outcomes drawn with thick lines: the observer sees the white disk as composed of two blurry spots of different sizes side by side, one light and one dark. That description matches the percepts reported by observers. This explanation was also supported by experimental results (Gori et al., 2010) showing that without the dynamical superimposition of the blurred-spot afterimage, the illusion disappears. The same mechanism could also explain the new motion illusion published by Takahashi, Niimi, and Watanabe (2010). These authors reported an illusory misperception of motion characterizing figures with blurry edges. This apparent motion could be caused by the superimposition of the afterimages on the physical stimuli like in the BLI.

In sum, this chapter reports a new illusory phenomenon called the BLI. This pattern presents a striking variation in size and brightness when observed dynamically. The luminance gradient seems to be the main prerequisite to trigger the illusory effects. The evidence discussed here indicates that the BLI emerges at retinal level, even if a cortical contribution cannot be excluded yet.

REFERENCES

Adelson, E. H. (1993). Perceptual organization and the judgment of brightness. *Science*, *262*(5142), 2042–2044.

Agostini, T., & Galmonte, A. (2002a). A new effect of luminance gradient on achromatic simultaneous contrast. *Psychonomic Bulletin and Review*, *9*(2), 264–269.

Agostini, T., & Galmonte, A. (2002b). Perceptual organization overcomes the effect of local surround in determining simultaneous lightness contrast. *Psychological Science, 13*(1), 88–92.

Agostini, T., & Proffitt, D. R. (1993). Perceptual organization evoked simultaneous lightness contrast. *Perception, 22*(3), 263–272.

Anstis, S., Gori, S., & Wehrhahn, C. (2007). Afterimages and the breathing light illusion. *Perception, 36*(5), 791–794.

Benary, W. (1924). Beobachtungen zu einem Experiment über Helligkeitskontrast. *Psychologische Forschung, 5*, 131–142. (Abridged English translation: The influence of form on brightness contrast. In W. Ellis [Ed.], *A source book of Gestalt psychology* [pp. 104–108]. London: Routledge & Kegan Paul, 1939)

Bressan, P. (2006). The place of white in a world of grays: A double-anchoring theory of lightness perception. *Psychological Review, 113*(3), 526–553.

Burgh, P. (1964). Peripheral viewing and simultaneous contrast. *Quarterly Journal of Experimental Psychology, 16*(3), 257–263.

Cohen, M. A., & Grossberg, S. (1984). Neural dynamics of brightness perception: Features, boundaries, diffusion, and resonance. *Perception & Psychophysics, 36*(5), 428–456.

Cornsweet, T. (1970). *Visual perception*. New York: Academic Press.

Economou, E., Zdravković, S., & Gilchrist, A. (2007). Anchoring versus spatial filtering accounts of simultaneous lightness contrast. *Journal of Vision, 7*(12), 2.1–2.15.

Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., . . . Economou, E. (1999). An anchoring theory of lightness perception. *Psychological Review, 106*(4), 795–834.

Gori, S., Giora, E., & Agostini, T. (2008). "Percept-percept coupling" in a case of illusory brightness variation: Some phenomenological considerations. *Teorie & Modelli*, *13*(2–3), 185–194.

Gori, S., Giora, E., & Agostini, T. (2010). Measuring the breathing light illusion by means of induced simultaneous contrast. *Perception, 39*(1), 5–12.

Gori, S., & Hamburger, K. (2006). A new motion illusion: The rotating-tilted-lines illusion. *Perception, 35*(6), 853–857.

Gori, S., & Spillmann, L. (2010). Detection vs. grouping thresholds for elements differing in spacing, size and luminance. An alternative approach towards the psychophysics of Gestalten. *Vision Research, 50*(12), 1194–1202.

Gori, S., & Stubbs, D. A. (2006). A new set of illusions: The dynamic luminance gradient illusion and the breathing light illusion. *Perception, 35*(11), 1573–1577.

Gori, S., & Yazdanbakhsh, A. (2008). The riddle of the rotating tilted lines illusion. *Perception, 37*(4), 631–635.

Heinemann, E. G. (1955). Simultaneous brightness induction as a function of inducing- and test-field luminances. *Journal of Experimental Psychology, 50*(2), 89–96.

Helmholtz, H. (1867). *Handbuch der physiologischen Optik*. Leipzig, Germany: Voss. (English translation: *Treatise on physiological optics*, edited by J. P. L. Southall. New York: Dover, 1962)

Hering, E. (1878). *Zur Lehre vom Lichtsinn*. Vienna: Gerald und Sohn. (English translation: *Outlines of a theory of the light sense*, edited by L. M. Hurvich & D. Jameson. Cambridge, MA: Harvard University Press, 1964)

Hess, C., & Pretori, H. (1894). Messende Untersuchungen über die Gesetzmässigkeit des simultanen Helligkeits-Contrastes. *Albrecht von Graef's Archiv für Ophthalmologie, 40*, 1–24. (English translation: H. Flock & J. H. Tenney [1970]. Quantitative investigation of the lawfulness of simultaneous brightness contrast. *Perceptual and Motor Skills, 31*(3), 947–969)

Jameson, D., & Hurvich, L. M. (1964). Theory of brightness and color contrast in human vision. *Vision Research, 4*(1), 135–154.

Kanizsa, G. (1954). Il gradiente marginale come fattore dell'aspetto fenomenico dei colori. *Archivio di Psicologia Neurologia e Psichiatria*, *15*(3), 251–264.

Kanizsa, G. (1957). Gradient marginal et perception chromatique. In E. Mayerson (Ed.), *Problèmes de la couleur* (pp. 107–113). Paris: SEVPEN.

Kennedy, J. M. (1976). Sun figure: An illusory diffuse contour resulting from an arrangement of dots. *Perception, 5*(4), 479–481.

Kingdom, F. A. A. (1997). Simultaneous contrast: The legacies of Hering and Helmholtz. *Perception, 26*(6), 673–677.

Kingdom, F. A. A. (2008). Perceiving light versus material. *Vision Research, 48*(20), 2090–2105.

Kingdom, F. A. A. (2011). Lightness, brightness and transparency: A quarter century of new ideas, captivating demonstrations and unrelenting controversy. *Vision Research, 51*(7), 652–673.

Kingdom, F., & Moulden, B. (1988). Border effects on brightness: A review of findings, models and issues. *Spatial Vision, 3*(4), 225–262.

Kozaki, A., & Noguchi, K. (1976). The relationship between perceived surface-lightness and perceived illumination. *Psychological Research, 39*(1), 1–16.

Leonards, U., Troscianko, T., Lazeyras, F., & Ibanez, V. (2005). Cortical distinction between the neural encoding of objects that appear to glow and those that do not. *Cognitive Brain Research, 24*(1), 173–176.

Lu, H., Zavagno, D., & Liu, Z. (2006). The glare effect does not give rise to a longer-lasting afterimage. *Perception, 35*(5), 701–707.

Mach, E. (1865). Über die Wirkung der räumlichen Vertheilung des Lichtreizes auf die Netzhaut. *Sitzungsberichte der kaiserlichen Akademie der Wissenschaften*, *Mathematisch-naturwissenschaftliche Classe*, *52*, 303–322. (English translation: *Mach bands: Quantitative studies on neural networks in the retina*, edited by F. Ratliff. San Francisco, CA: Holden-Day, 1965)

MacLeod, R. (1947). The effects of "artificial penumbrae" on the brightness of included areas. In *Miscellanea Psychologica Albert Michotte* (pp. 138–154). Louvain and Paris: Éditions de l'Institut Supérieur de Philosophie and Vrin.

Petrov, Y. A., & Popple, A. V. (2002). Effects of negative afterimages in visual illusions. *Journal of the Optical Society of America A, 19*(6), 1107–1111.

Pinna, B., Brelstaff, G., & Spillmann, L. (2001). Surface color from boundaries: A new "watercolor" illusion. *Vision Research, 41*(20), 2669–2676.

Ross, H. E., & Plug, C. (1998). The history of size constancy and size illusions. In V. Walsh & J. Kulikowski (Eds.), *Perceptual constancy* (pp. 499–528). Cambridge, MA: Cambridge University Press.

Spillmann, L. (1997). Colour in a larger perspective: The rebirth of Gestalt psychology. *Perception, 26*(11), 1341–1352.

Spillmann, L. (2009). Phenomenology and neurophysiological correlations: Two approaches to perception research. *Vision Research, 49*(12), 1507–1521.

Spillmann, L., Hardy, J., Delahunt, P., Pinna, B., & Werner, J. S. (2010). Brightness enhancement seen through a tube. *Perception, 39*(11), 1504–1513.

Spillmann, L., & Werner, J. S. (1996). Long-range interactions in visual perception. *Trends in Neuroscience, 19*(10), 428–434.

Takahashi, K., Niimi, R., & Watanabe, K. (2010). Illusory motion induced by blurred red-blue edges. *Perception, 39*(12), 1678–1680.

Todorović, D. (2006). Lightness, illumination, and gradients. *Spatial Vision, 19*(2–4), 219–261.

Troncoso, X. G., Macknik, S. L., & Martinez-Conde, S. (2005). Novel visual illusions related to Vasarely's "nested squares" show that corner salience varies with corner angle. *Perception, 34*(4), 409–420.

Vasarely, V. (1970). *Vasarely II.* Neuchâtel, Switzerland: Éditions du Griffon.

Wade, N. J. (1996). Descriptions of visual phenomena from Aristotle to Wheatstone. *Perception, 25*(10), 1137–1175.

Wallach, H. (1948). Brightness constancy and the nature of achromatic colors. *Journal of Experimental Psychology, 38*(3), 310–324.

Yazdanbakhsh, A., & Gori, S. (2008). A new psychophysical estimation of the receptive field size. *Neuroscience Letters, 438*(2), 246–251.

Zavagno, D. (1999). Some new luminance-gradient effects. *Perception, 28*(7), 835–838.

Zavagno, D., & Massironi, M. (2006). Colours in black and white: The depiction of lightness and brightness in achromatic engravings before the invention of photography. *Perception, 35*(1), 91–100.

Zdravković, S., Economou, E., & Gilchrist, A. (2012). Grouping illumination frameworks. *Journal of Experimental Psychology: Human Perception and Performance, 38*(3), 776–784.

Chapter 48
Large Shift in Brightness Induced by Motion in Context

Sang Wook Hong

INTRODUCTION

Visual scenes we regularly encounter in everyday life are naturally cluttered, in that many objects are scattered all over the place and are dynamically moving in many directions. Is the appearance of an object, in terms of brightness and color, always constant, irrespective of whether it is moving or stationary and whether it is presented alone or with other objects? A novel brightness illusion, introduced by Hong and Kang (2013), suggests that a stationary object can appear largely different when it is viewed with other objects moving around it.

In a laboratory setting, two pairs of gray dots (7.5 cd/m^2) were presented on a uniform white background. At the beginning of the display, all four dots were stationary and physically identical so that they appeared the same. When one of the dots in a pair started orbiting around the other stationary dot of the pair, the brightness of both the stationary and the moving dots changed instantaneously (Fig. III.48-1a; see Video III.48-1 and Video III.48-2 on the Companion Website). Laboratory measurements show that the stationary dot appeared much brighter (matched luminance was about 23 cd/m^2) and the moving dot appeared slightly darker (matched luminance was about 5 cd/m^2) than the brightness of the dots before the motion onset (Fig. III.48-1b). This illusory brightness change was robust and strong, so it can be observed without sophisticated settings used in laboratories. One can perceive a large shift in brightness immediately after the onset of motion, and the illusory brightness shift is perceived continuously as long as the motion lasts. This novel brightness illusion is distinguished from any other brightness induction by spatially and temporally varying background context. First, brightness of a stationary object changes simply due to the movement of a physically identical one without any change in overall light level in the surrounding context. Second, the brightness shift occurs when the distance between the stationary and the moving objects is as far as 5° of visual angle. This large spatial range of interaction between the target stimulus and the inducing background has never been observed previously. Generally speaking, spatial distance between the target object and the inducing surround is one of the critical factors in brightness illusions (Hong & Shevell, 2004b; Rudd & Zemach, 2004).

POSSIBLE NEURAL MECHANISMS

The human brain dynamically regulates sensory processing. This capacity of our brain allows us to discriminate or detect subtle differences between visual inputs in our ever-changing environment. For example, when we leave a very dark room and enter a very bright and sunny space, our visual system quickly adapts. Our pupils shrink and receptors in the retina adjust their light absorption, and, consequently, neural responses along the visual processing stream change accordingly. Along with this light adaptation example, various types of adaptive sensitivity regulation also occur at multiple stages of visual processing. This incredible sensitivity regulation allows us to function properly in dynamically changing environments (i.e., driving in and out of a tunnel).

Perception of color and brightness of an isolated object is determined mainly by physical properties of light (e.g., intensity and spectral power distribution) reflected from the object. When the object is integrated into a complex scene, the appearance of the object can change dramatically. Changes in perceived color and brightness often result from a type of sensitivity regulation called adaptation. Adaptation to a specific environment changes neural responses to a visual stimulus embedded in the environment and consequently changes the appearance of the stimulus. For example, color appearance depends on spatial and temporal adaptation to the average color (chromatic adaptation) and to the variation in color (chromatic contrast adaptation) in the visual scene. The gain control mechanism, an adaptive sensitivity regulation mechanism, in receptoral and postreceptoral levels provides a decent explanation for both chromatic adaptation (Chichilnisky & Wandell, 1995; Shevell, 1982; von Kries, 1905/1970; Walraven, 1976; Wei & Shevell, 1995) and chromatic contrast adaptation (Barnes, Wei, & Shevell, 1999; Singer & D'Zmura, 1994, 1995; Webster & Mollon, 1991, 1994).

Although brightness is sometimes considered one of three dimensions of color (along with hue and saturation), perception of an object's (or surface's) brightness is not simple. Decades of research reveals that a subjective experience of brightness involves complicated neural processing that takes into account illumination (Schirillo & Shevell, 1997), surface reflectance (Gilchrist et al., 1999), and the spatial and temporal relation between an object and its context (Adelson,

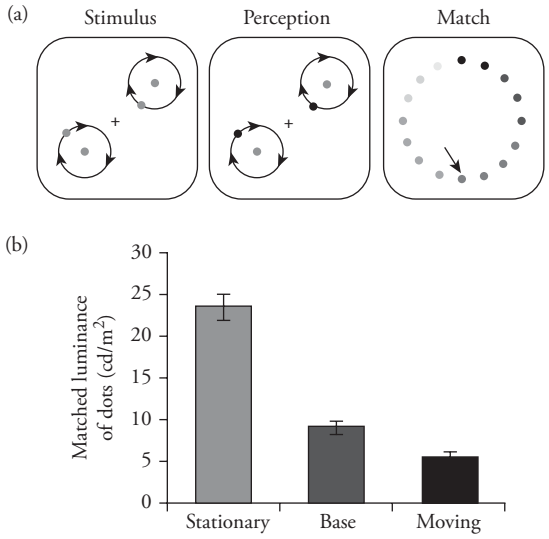
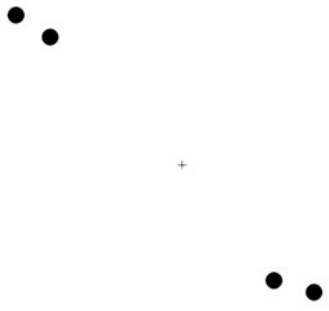
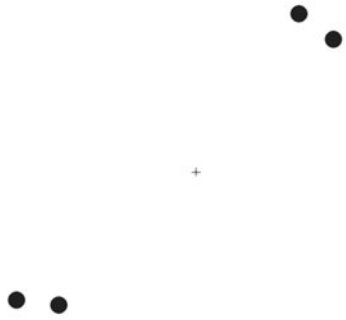


Figure III.48-1. (a) Schematic illustration of brightness shift due to motion in context. Two pairs of identical gray dots are presented, and one dot in each pair is orbiting around a stationary dot (left). The stationary dots appear brighter due to motion in the surrounding context (middle). After the stimulus is turned off, an array of objects with luminance or color patches will appear. Observers are instructed to choose the one closest to the perceived appearance of the stationary target (right). (b) Brightness matching results from the memory-based matching task, adopted from Hong and Kang (2013). The "base" represents the appearance of dots when all four dots are stationary.

1993; De Valois, Webster, De Valois, & Lingelbach, 1986; Gilchrist, 1977; Schirillo, 1999; Todorović, 1997). Theories of brightness posit that computation of an object's brightness involves both contrast between the object and its adjacent surround and also between the object and remote surround (Hong & Shevell, 2004a, 2004b; Rudd & Arrington, 2001; Shapley & Reid, 1985; Shevell, Holliday, & Whittle, 1992). Since processing of contrast is closely tied to brightness, the contrast gain control has been also considered to understand brightness. Using an oriented difference of Gaussian model,



Video III.48-1. At the beginning of the video, a fixation cross appears in the middle of the display window. To observe the illusion best, fixate your eyes on the cross throughout the video (although slight eye movement does not abolish the illusion). Two pairs of black dots (all stationary) are presented for 1 s, then one of the dots in each pair starts orbiting around the other stationary dots. The brightness of the stationary dots appears gray right after the motion onset.



Video III.48-2. This is the same as Video III.48-1, except here all dots are dark gray initially instead of black. This demonstrates that the appearance of a moving dot is also different from the appearance of the same dot before the motion onset. The dot appears darker when it starts moving.

Blakeslee and McCourt (1999, 2004) show that the brightness of a surface embedded in a complex surround can be explained by a contrast gain control mechanism operating with response normalization.

It is not clear, however, whether the contrast gain control can explain the large shift in brightness induced by motion in context. Although several characteristics of the brightness shift, such as long-range interaction between the stationary and the moving dots and contrast polarity dependency (Hong & Kang, 2013), are consistent with the predictions based on the contrast gain control model, several questions remain. First, how does simple motion of a physically identical object change neural responses to the stationary object? It is not known whether motion signal can trigger contrast gain control over the whole visual space. Second, does the same motion in context directly change apparent contrast as well as brightness? Without the answers to these questions, it will remain hard to completely understand the mechanisms mediating the brightness illusion reported here.

IMPLICATIONS FOR OTHER VISUAL ILLUSIONS

As mentioned earlier, visual scenes in a natural environment are usually cluttered with many objects and are also dynamically changing. Competition among visual objects for conscious perceptual experience is a natural consequence of the limited capacity of the human visual system, which can be easily overwhelmed by the information surging into the eyes at every moment. Visual competition often results in complete disappearance of one of the competing inputs from visual awareness, and this is compellingly demonstrated by variety of phenomena including motion-induced blindness (MIB) in which a visually salient stimulus keeps disappearing and reappearing when it is surrounded by a dynamically moving background (Bonneh, Cooperman, & Sagi, 2001); binocular and monocular rivalry, in which only one of two dissimilar images presented binocularly or monocularly is consciously perceived (Andrews & Pruves, 1997;

Blake & Logothetis, 2002); and bistable perception (such as Necker cube and vase/face illusion), in which two possible representations of one physical stimulus compete with one another for perceptual dominance. Vision science until now has been devoted to understanding the mechanisms that determine whether or not an object is consciously perceived, but it has completely ignored perceptual quality, such as the appearance of an object, resulting from visual competitions.

The significance of the brightness illusion introduced here is that it sheds light on a different consequence of visual competition other than complete disappearance of an object from visual awareness. For example, the brightness shift induced by motion is similar to MIB. Both visual phenomena are induced by the interaction between stationary objects and moving objects in the surroundings, but one results in a shift in brightness and the other results in a complete disappearance from awareness. The similarity between MIB and the brightness shift due to motion in context suggests that visual competition between the stationary and the moving objects and its consequent sensitivity regulation can be the low-level sensory mechanism of MIB rather than prolonged adaptation to the stationary object (Bonneh et al., 2001; Gorea & Caetta, 2009). Other mechanisms explaining MIB, such as disruption of attention and motion streak, then operate further to modulate conscious experience of the stationary objects (Bonneh et al., 2001; Wallis & Arnold, 2009).

Despite all the advancements in understanding the perceptual experience resulting from visual competition, only a few studies have shown that visual competition between objects may change the appearance of objects. For example, Carrasco, Ling, and Read (2004) show that directing attention to one of two objects makes the perceived contrast of the object higher than the other, nonattended object. In this case, the target objects are competing against each other for attentional resources, and the "loser" of the competition appears less salient than the "winner," which is consistent with the brightness illusion introduced here. Thus by probing the neural mechanisms, this brightness illusion will provide clues to answer a long-lasting question in vision science: How does the sensory processing of the human visual system deal with multiple objects in a cluttered visual scene aside from ignoring task-irrelevant or unimportant objects through attentional control?

REFERENCES

Adelson, E. H. (1993). Perceptual organization and the judgment of brightness. *Science*, *262*, 2042–2044.

Andrews, T. J., & Purves, D. (1997). Similarities in normal and binocularly rivalrous viewing. *Proceedings of the National Academy of Sciences USA*, *94*, 9905–9908.

Barnes, C. S., Wei, J., & Shevell, S. K. (1999). Chromatic induction with remote chromatic contrast varied in magnitude, spatial frequency, and chromaticity. *Vision Research*, *39*, 3561–3574.

Blake, R., & Logothetis, N. (2002). Visual competition. *Nature Reviews Neuroscience*, *3*, 13–21.

Blakeslee, B., & McCourt, M. E. (1999). A multiscale spatial filtering account of the White effect, simultaneous brightness contrast and grating induction. *Vision Research*, *39*, 4361–4377.

Blakeslee, B., & McCourt, M. E. (2004). A unified theory of brightness contrast and assimilation incorporating oriented multiscale spatial filtering and contrast normalization. *Vision Research*, *44*, 2483–2503.

Bonneh, Y. S., Cooperman, A., & Sagi, D. (2001). Motion-induced blindness in normal observers. *Nature*, *411*, 798–801.

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, *7*, 308–313.

Chichilnisky, E. J., & Wandell, B. (1995). Photoreceptor sensitivity changes explain color appearance shifts induced by large uniform backgrounds in dichoptic matching. *Vision Research*, *35*, 239–254.

De Valois, R. L., Webster, M. A., De Valois, K. K., & Lingelbach, B. (1986). Temporal properties of brightness and color induction. *Vision Research*, *26*, 887–897.

Gilchrist, A. L. (1977). Perceived lightness depends on perceived spatial arrangement. *Science*, *195*, 185–187.

Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., … Economou, E. (1999). An anchoring theory of lightness perception. *Psychological Review*, *106*, 795–834.

Gorea, A., & Caetta, F. (2009). Adaptation and prolonged inhibition as a main cause of motion-induced blindness. *Journal of Vision*, *9*(6), 16, 1–17.

Hong, S. W., & Kang, M.-S. (2013). Perceptual consequence of normalization revealed by a novel brightness induction. *Vision Research*, *91*, 78–83.

Hong, S. W., & Shevell, S. K. (2004a). Brightness contrast and assimilation from patterned inducing backgrounds. *Vision Research*, *44*, 35–43.

Hong, S. W., & Shevell, S. K. (2004b). Brightness induction: Unequal spatial integration with increments and decrements. *Visual Neuroscience*, *21*, 353–357.

Rudd, M. E., & Arrington, K. F. (2001). Darkness filling-in: A neural model of darkness induction. *Vision Research*, *41*, 3649–3662.

Rudd, M. E., & Zemach, I. K. (2004). Quantitative properties of achromatic color induction: An edge integration analysis. *Vision Research*, *44*, 971–981.

Schirillo, J. (1999). Surround articulation. I. Brightness judgments. *Journal of the Optical Society America A*, *16*, 793–803.

Schirillo, J. A., & Shevell, S. K. (1997). An account of brightness in complex scenes based on inferred illumination. *Perception*, *26*, 507–518.

Shapley, R., & Reid, R. C. (1985). Contrast and assimilation in the perception of brightness. *Proceedings of the National Academy of Sciences USA*, *82*, 5983–5986.

Shevell, S. K. (1982). Color perception under chromatic adaptation: Equilibrium yellow and long-wavelength adaptation. *Vision Research*, *22*, 279–292.

Shevell, S. K., Holliday, I., & Whittle, P. (1992). Two separate neural mechanisms of brightness induction. *Vision Research*, *32*, 2331–2340.

Singer, B., & D'Zmura, M. (1994). Color contrast induction. *Vision Research*, *34*, 3111–3126.

Singer, B., & D'Zmura, M. (1995). Contrast gain control: A bilinear model for chromatic selectivity. *Journal of the Optical Society America A*, *12*, 667–685.

Todorović, D. (1997). Lightness and junctions. *Perception*, *26*, 379–394.

von Kries, J. (1970). Influence of adaptation on the effects produced by luminous stimuli. In D. L. MacAdam (Ed.), *Sources of color science* (pp. 120–126). Cambridge, MA: MIT Press. (Original work published 1905)

Wallis, T. S. A., & Arnold, D. H. (2009). Motion-induced blindness and motion streak suppression. *Current Biology*, *19*, 325–329.

Walraven, J. (1976). Discounting the background—The missing link in the explanation of chromatic induction. *Vision Research*, *16*, 289–295.

Webster, M. A., & Mollon, J. D. (1991). Changes in colour appearance following post-receptoral adaptation. *Nature*, *349*, 235–238.

Webster, M. A., & Mollon, J. D. (1994). The influence of contrast adaptation on color appearance. *Vision Research*, *34*, 1993–2020.

Wei, J., & Shevell, S. K. (1995). Color appearance under chromatic adaptation varied along theoretically significant axes in color space. *Journal of the Optical Society of America A*, *12*, 36–46.

Chapter 49
The Chromatic Mach Card

Anya Hurlbert

INTRODUCTION

In his *Analysis of Sensations*, the eponymous Ernst Mach (1886/1959) used his card to illustrate the intrinsic "intelligence" of "the eye" and the "fixed habit" by which it has learned to see the relationship between illumination and depth. He observed that when a rectangular card, folded lengthwise, is placed with its edge facing outward and the light falls from the left, one scarcely perceives the physical difference in brightness of the two sides. When the viewer closes one eye and forces upon him or herself the impression that the card is concave, with the bent edge facing inward instead of outward, the difference in brightness between the two sides becomes profound, as if they were painted two deeply different shades of grey (Fig. III.49-1).

The effect demonstrates that objects' shape and surface properties are intertwined in visual perception—the lightness, or perceived surface reflectance, of the card is dependent on its perceived shape. Mach (1886/1959) argued that the purpose of seeing is to define and locate objects in space, not to draw out light and shade. Thus although the real luminance difference between the two sides of the card feeds into an unconscious estimation of their difference in depth and spatial orientation, it is then discarded from conscious perception of the card as a unitary, whole object. He concluded that "[d]ifferences of brightness are partly transformed into differences of depth, and themselves become weaker in the process" (p. 211).

The fact that this process was clearly inaccessible to consciousness implied to Mach that it must not occur in the cerebrum, where mental imagery could clearly be distinguished from direct sensations. It must therefore occur in the eye itself. Indeed, because the effect wholly depends on the assumption of a fixed direction of illumination, Mach proposed that the eyeball itself, being translucent, directly sensed the impinging light. Hence, if the eye saw the card as concave—a perfectly plausible shape given the limited depth information from one eye—it must also then see the brightness difference as a real difference of surface reflectance, because the illumination direction remained firmly fixed.

Thus Mach localized the perceptual interdependence of scene and object attributes—illumination, reflectance, and depth—to very early stages of visual processing. One hundred years later, in the late 20th century, the emphasis was instead on the functional modularity of vision (DeYoe, Felleman, Van Essen, & McClendon, 1994; Livingstone & Hubel, 1988). The prevalent view was that the eye performed only the basic spatial and temporal decomposition of the light signal, while at higher levels, distinct cortical areas or subregions thereof were dedicated to processing distinct attributes such as form, color, and motion (Hubel & Livingstone, 1987). Where and how in the visual system these attributes were integrated in holistic object representations was (and still is) a subject of investigation and debate (Cant & Goodale, 2007; Cavina-Pratesi, Kentridge, Heywood, & Milner, 2010; DiCarlo, Zoccolan, & Rust, 2012).

The question of where object color itself is represented is tied up with the question of where and how color constancy is achieved. For most natural objects, color is integral to

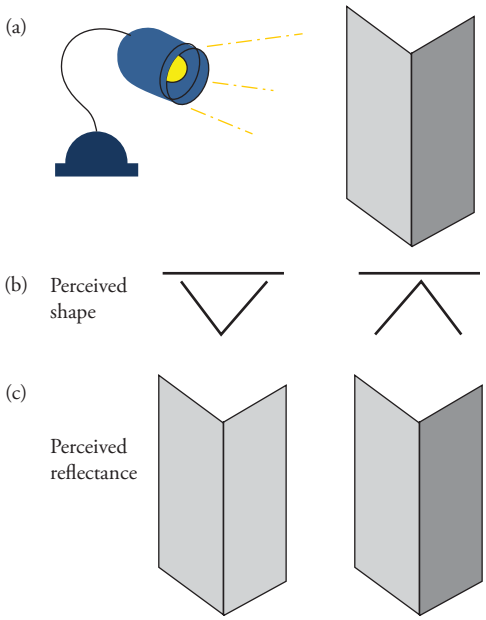


Figure III.49-1. The classical Mach card. (a) The light source shines on the left side of a convex folded card, made from uniform white paper. (b) and (c) When the card appears veridically convex (left column), the two sides are perceived to have the same surface reflectance. When the card appears concave (right column), its right side is perceived to have darker reflectance than the left. Figure credit: Hurlbert.

object identity—the redness of a tomato, the greenness of grass, the flesh color of human skin—yet the spectral light signal from objects is highly variable. The spectrum of light reflected from an object depends on both its intrinsic surface reflectance and the extrinsic illumination spectrum, which may change dramatically from moment to moment and place to place. The extraction of constant color from this variable light signal would seem essential for the representation of object color. Most evidence points to a relatively late stage of visual processing as the site of color constancy, which places color representations at similarly late stages. Indeed, Crick and Koch (1995) argued against primary visual cortex as the site for consciousness, specifically because neurons there do not display the properties necessary for color constancy, and color constancy is a bedrock of our conscious experience.

Although Mach did not explicitly refer to color, and the 19th-century visiting card he used to demonstrate his effect typically would have been white, nonetheless his effect suggests that the apparent intrinsic surface reflectance properties of objects—and, hence, their colors—must be inextricably linked to the perception of three-dimensional (3D) shape and scene configuration.

In the late 1990s, Dan Kersten, Marina Bloj, and I revived the Mach card but this time in color. The two sides of the chromatic Mach card are painted different colors, one side magenta and the other white. When the card is open, the magenta side casts its reflections on the white side, causing it, in turn, to reflect pinkish light to the eye. Yet viewers perceive the white side to be white, or nearly white, unconsciously and accurately compensating for the effect of the interreflections to recover the "true" color of the white side. But when viewers see exactly the same card as if it has been inverted in depth, so that the central fold appears to protrude outward toward the observer, the white side appears a deep pink (Fig. III.49-2). Its color flips when the perceived shape flips. Our explanation for the effect was that the seeing brain can no longer explain away the pink glow from the white side as caused by reflections from the other side, because the two sides no longer face each other. Therefore, the brain concludes that the white side must itself be painted pink.

Like Mach's original card, the chromatic Mach card shows that the perception of object color is linked to the perception of object shape and illumination. But there are key differences between the two phenomena and their explanations.

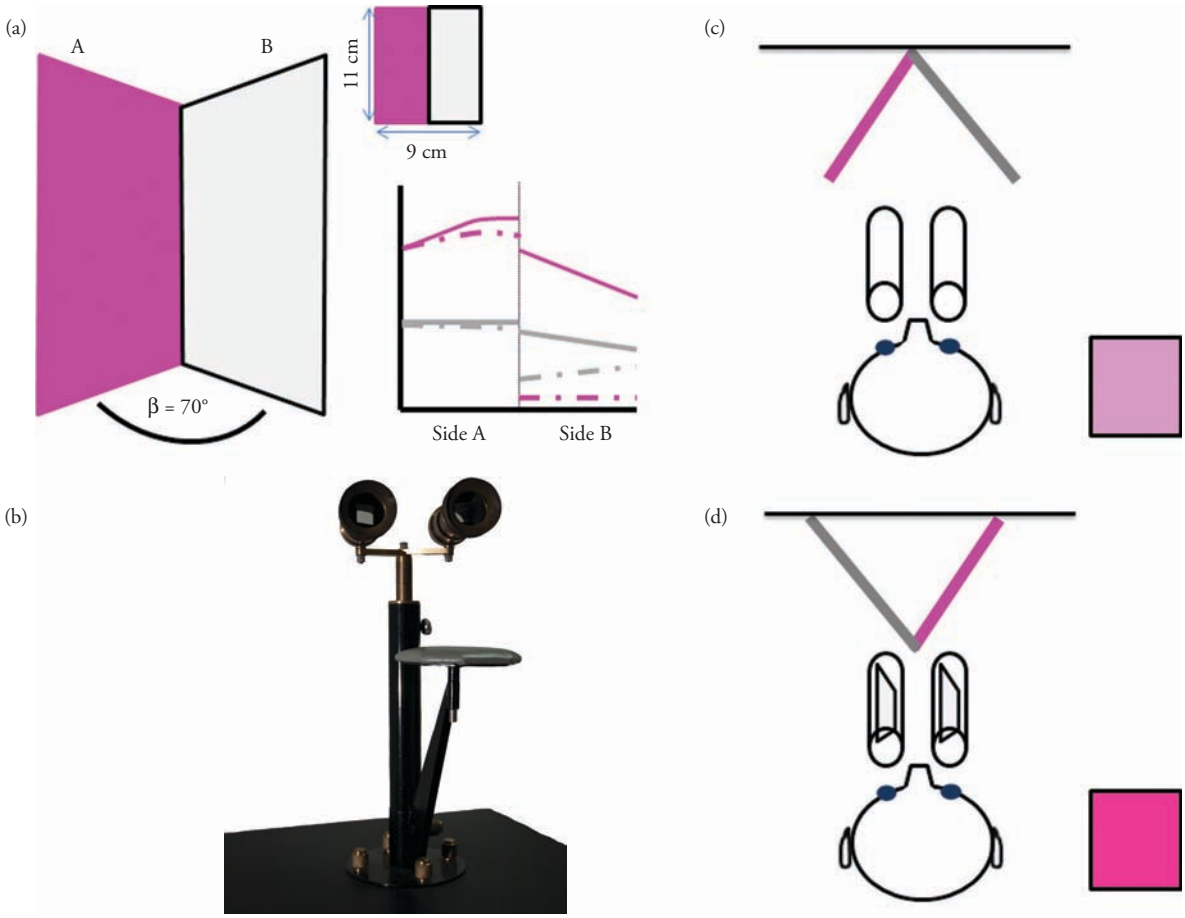


Figure III.49-2. The chromatic Mach card. (a) Schematic drawing of card used in experimental demonstration of the effect (Bloj et al., 1999). Opening angle of card: 70°. Inset (top): Planar view of card with actual dimensions. Inset (lower): Chroma (pink lines) and lightness (grey lines) sampled along horizontal scan line. Solid lines: With mutual reflections. Dashed lines: Without mutual reflections. (b) Pseudoscope, with Dove prism-fitted cylinders. (c) The card viewed through empty cylinders appears concave, and side B is matched to a pinkish-white patch (low chroma). (d) The card viewed through Dove prisms appears depth-inverted and side-reversed, and side B (now on left) is matched to a deep-pink patch (high chroma). Figure credit: Hurlbert.

In the classical effect, Mach's explanation requires that the viewer continues to perceive the illumination as coming from the left; its direction does not reverse when the shape reverses. Thus when the darker side appears to be turned toward the illumination, the "eye" must conclude that it is actually a darker grey than the bright side. Beck (1965) demonstrated that viewers' assumptions about the number and positions of the light sources on the card did indeed influence their lightness matches to the two sides.

In contrast, the explanation for the chromatic Mach card effect is that the viewer is uncertain of the external illumination direction but intrinsically certain of the physics of mutual illumination. Mutual illumination is rife in the natural environment, created by surfaces reflecting light onto each other. Because mutual illumination depends strongly on the 3D geometry of the surfaces involved, such interreflections may provide important cues to object shape, alongside the shading cues to shape from the direct illumination, which are exploited in classic shape-from-shading computer vision algorithms (Horn, 1977; Ikeuchi & Horn, 1981). The chromatic Mach card demonstrates that in human vision perceived object shape dictates perceived mutual illumination. When the card is convex, pointing outward, mutual illumination between the two sides is impossible. The visual system understands this relationship and uses it to guide interpretation of an otherwise ambiguous image.

The empirical measurements we made of the chromatic Mach card effect reflect this interpretive process (Bloj, Kersten, & Hurlbert, 1999). A concave magenta–white card with trapezoidal sides was placed on the back wall of a black box and illuminated from the right by a hidden incandescent light bulb (Fig. III.49-2). Observers viewed the card from the open front end of the box through binocular viewing tubes. In the "roof" condition, the viewing cylinders contained Dove prisms (making a pseudoscope), which invert the image in each eye from left to right and thereby reversed the binocular disparities of the image, thus causing the card to appear convex. In the "corner" condition, the cylinders were identical but empty, and the card appeared in its actual concave configuration. Observers were asked to pick the color that best matched the color of the left ("roof") or right ("corner") side of the card—a straightforward color appearance task. They were not asked to judge the physical surface reflectance properties of either side. Yet the range of matches that observers chose were very well predicted by a probabilistic model that assumes that the visual system is making its best attempt to recover the surface reflectance of the white side in the face of uncertain environmental variations.

The model, based on Bayesian inference, calculates how an ideal observer would behave if asked to select the matching chip for the white side under the two conditions. The ideal observer understands the physics of light reflection and has a priori knowledge of the primary light source spectrum, the surface reflectances of all the possible matching chips, and the surface reflectance of the magenta side. The ideal observer is not particularly privileged in possessing this information, simply experienced in the ways of the world: tungsten bulbs are very common indoor light sources (or at least were, at the time of this experiment); magenta paints have characteristically limited properties; both types of spectra may be reconstructed from small sets of basis functions.

But the ideal observer is ignorant of the primary light source direction—in part because it has been deliberately concealed in this experiment and in part, we suggest, because the visual system has evolved not to waste too much energy pinning down such a naturally mutable property. The angle of the sun may change, but this should not affect one's assessment of an object's redness, if, say, one wants to find ripe-red berries. Thus the visual system should attempt to find the most likely surface reflectance, given accidental variations in the light source direction. The model computes the difference between the actual light signal from the white side and that predicted for a given matching chip for each possible illumination direction and converts the sum of these differences into a probability; the smaller the differences overall, the more likely the particular matching chip. The computation relies on the visual system instantiating the formula for light reflection, which includes terms both for the direct illumination from the external source (with varying direction), and indirect mutual illumination. The probability distributions for the matching chip are very different for the concave card, when both terms contribute, and the convex card, when only the direct illumination exists. And the computed probability distributions very neatly predict the range and number of matches that observers choose (Fig. III.49-3).

Mutual illumination does not explain the classical Mach card effect, and the classical Mach card effect does not explain the chromatic Mach card effect. The main effect of mutual illumination in the concave achromatic card is to increase the brightness at the crease, where "white" light bounces back and forth between the two sides. As Gilchrist and Jacobsen (1984) showed, such bright strips in concave corners help viewers to distinguish an all-white room under very dim illumination from a nearly black room under very bright light; in the latter room, mutual illumination is almost nonexistent. But the brightening at the crease also lessens the overall brightness gradient across the card, and the reduction in the brightness gradient cue might work against accurate estimation of the slant, in turn weakening the crucial contribution of shape perception to the effect. In computer vision, these alterations to the brightness gradient cause classic shape-from-shading algorithms—based on direct illumination effects only—to fail unless additional constraints are enforced (Forsyth, 2011; Forsyth & Zisserman, 1990; Nayar, Ikeuchi, & Kanade, 1991). Whether the human visual system uses all of the information—from direct illumination and mutual illumination—available in brightness gradients to recover shape remains an open question; recent experiments suggest that humans require training to use the information conveyed by combined color and brightness gradients (Harding, Harris, & Bloj, 2012). But whether or not the effects of achromatic mutual illumination contribute to shape perception, they are not sufficient to induce the Mach card effect in the absence of knowledge of the light source direction and without perspective cues to shape, and neither do they make a difference to the predicted lightness matches when incorporated in a Bayesian model (Bloj & Hurlbert, 2002). Mutual illumination plays no role in the classical Mach card effect.

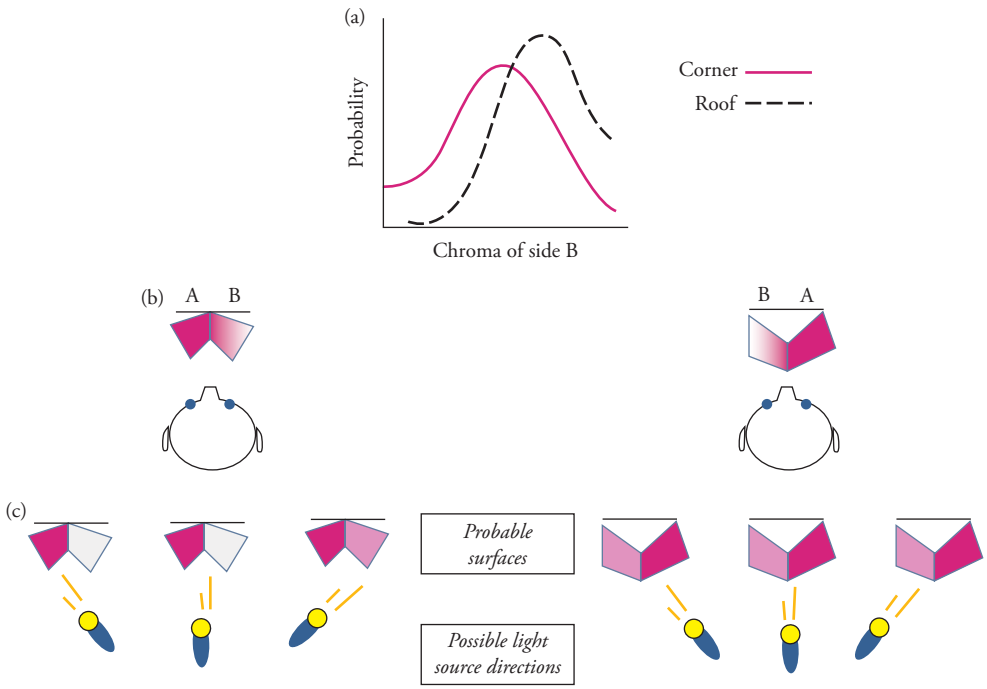


Figure III.49-3. The Bayesian model. (a) Predicted probability curves for chroma of side B, assuming known light source spectrum and surface reflectances of side A and all possible matching chips but unknown light-source direction. Dashed line: Perceived convex card (roof). Solid line: Perceived concave card (corner). (b) Given the image, the model of the ideal observer computes the probability of side B being a particular surface reflectance, taken over all possible light source directions. (c) Cartoon of the most probable surface reflectance of side B for particular light source directions for the corner and roof shapes. Figure credit: Hurlbert.

Conversely, neither does the classical Mach card effect explain the chromatic effect. If the viewer were able to perceive correctly the illumination as coming from the right, he or she would see the white side of the card as more strongly illuminated in its concave form (corner) and therefore match it to inherently darker shades of grey than in the convex (roof) form. But this difference does not occur. The perceptual flip in the chromatic Mach card is not a lightness effect. Indeed, since the chromatic effects of mutual illumination are more perceptually striking than the achromatic ones, it is not surprising that the chromatic effects of mutual illumination entirely explain the chromatic Mach card effect. It is plausible that, in human vision, the information conveyed by mutual illumination effects is used only to estimate coarse depth and spatial configuration, not to recover depth or 3D shape on a fine scale. In line with this notion, Madison, Thompson, Kersten, Shirley, and Smits (2001) demonstrated that the presence of mutual illumination was sufficient to reveal surface contact, a rudimentary depth feature.

KEY PARAMETERS

The chromatic Mach card effect works with simple card, deeply colored on one side and white on the other, and by simply shutting one eye, or using an act of will, to invert the shape. But to maximize the effect for the quantitative demonstration, we optimized the color, minimized the cues to the light source direction, and streamlined the shape inversion. To eliminate cues to the external illumination direction, we removed cast shadows by placing the card against a black felt background. To ease the shape inversion, we removed perspective cues to the card's actual shape by making its sides trapezoidal. We used pseudoscopic inversion of the binocular disparities to enhance the shape, inspired by Sasha Logvinenko's use of it (Logvinenko & Menshikova, 1994).

The hue of the colored side is not a crucial parameter: Dan Kersten and I originally made a yellow-white card (playing at the European Conference on Visual Perception in 1995) and immediately obtained the effect (I had been studying mutual illumination and color constancy using pleated Mondrians; see Fig. III.49-4; Bloj & Hurlbert 1995; Hurlbert, 1998a) and had just reduced the stimulus to a two-surface pleat, while Dan had been thinking about the effects of contour on reflectance perception (Knill & Kersten, 1991). Trial and error in the lab led us to a very bright magenta paint, which under the yellowish-orange tungsten light source appeared a very saturated red (CIE *xy* values of 0.662, 0.312). Laid flat under the light source, without mutual illumination, the white paper had CIE *xy* values of 0.509, 0.412 but would appear white because of chromatic adaptation; with mutual illumination from the magenta paper, its average CIE *xy* value was 0.636, 0.328, a more saturated red, yet it still appeared nearly white—the crux of the chromatic Mach card effect.

These CIE *xy* values would fall outside of the gamut of most typical cathode ray tube CRT displays, and thus the card cannot be reproduced exactly there. But the effect is difficult to achieve with computer simulations (Fig. III.49-4) primarily because of the difficulty of securing the 3D shape in the absence of surface features and contours: even with

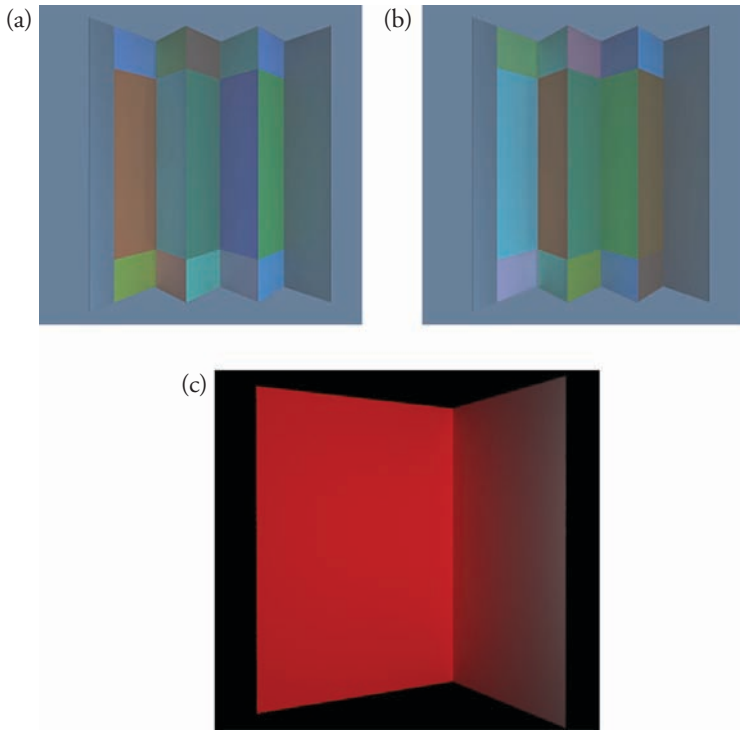


Figure III.49-4. Computer graphics simulations of mutual reflections. Pleated Mondrians: The central greenish patch appears darker and more turquoise in (a), where the neighboring reflecting patch is dark blue, than in (b), where the reflecting patch is bright green. (c) Simulation of chromatic Mach card. All images are made with a full radiosity algorithm; each distinct "paper" is assigned a uniform surface reflectance, and a single light source illuminates the scene. Figure credit: Hurlbert.

stereoscopic rendering and viewing, the 3D "corner" and "roof" shapes are frangible. Generally, the effect demonstrates that real life is better than what we can produce on a 2D screen (Hurlbert, 1998b).

CONCLUSIONS

The chromatic Mach card effect drives home the truth that color is a construct of the brain, not a fixed property of objects "out there." The color of a physically constant card changes in a somersault of perception. The effect also demonstrates that over the course of evolution our brains have learned much about the physics of light reflection—how light is bounced from one surface onto another, how the color of daylight changes from morning to evening—and that it has stored this knowledge in the automatic workings of the visual system. The color changes only because the physics of the scene appears also to have flipped. Thus, much as Mach (1886/1959) argued, the visual system incorporates an unconscious understanding: it is endowed with "its own purposive habits, its own peculiar memory, one might almost say its own intelligence" (p. 202).

REFERENCES

Beck, J. (1965). Apparent spatial position and the perception of lightness. *Journal of Experimental Psychology*, *69*, 170–179.

Bloj, M. G., & Hurlbert, A. C. (1995). Does mutual illumination improve human colour constancy? *Investigative Ophthalmology and Visual Science*, *36*(4), S639.

Bloj, M. G., & Hurlbert, A. C. (2002). An empirical study of the traditional Mach card effect. *Perception*, *31*, 233–246.

Bloj, M. G., Kersten, D., & Hurlbert, A. C. (1999). Perception of three-dimensional shape influences colour perception through mutual illumination. *Nature, 402*(6764), 877–879.

Cant, J. S., & Goodale, M. A. (2007). Attention to form or surface properties modulates different regions of human occipitotemporal cortex. *Cerebral Cortex*, *17*(3), 713–731.

Cavina-Pratesi, C., Kentridge, R. W., Heywood, C. A., & Milner, A. D. (2010). Separate processing of texture and form in the ventral stream: Evidence from fMRI and visual agnosia. *Cerebral Cortex*, *20*(2), 433–446.

Crick, F., & Koch, C. (1995). Are we aware of neural activity in primary visual-cortex? *Nature*, *375*(6527), 121–123.

DeYoe, E. A., Felleman, D. J., Van Essen, D. C., & McClendon, E. (1994). Multiple processing streams in occipitotemporal visual-cortex. *Nature*, *371*(6493), 151–154.

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How does the brain solve visual object recognition? *Neuron*, *73*(3), 415–434.

Forsyth, D. A. (2011). Variable-source shading analysis. *International Journal of Computer Vision*, *91*, 280–302.

Forsyth, D. A., & Zisserman, A. (1990). Shape from shading in the light of mutual illumination. *Image and Vision Computing*, *8*(1), 42–49.

Gilchrist, A., & Jacobsen, A. (1984). Perception of lightness and illumination in a world of one reflectance. *Perception*, *13*, 5–19.

Harding, G., Harris, J. M., & Bloj, M. (2012). Learning to use illumination gradients as an unambiguous cue to three dimensional shape. *PLoS ONE*, *7*(4), e35950.

Horn, B. K. P. (1977). Understanding image intensities. *Artificial Intelligence*, *8*(2), 201–231.

Hubel, D. H., & Livingstone, M. S. (1987). Segregation of form, color, and stereopsis in primate area-18. *Journal of Neuroscience*, *7*(11), 3378–3415.

Hurlbert, A. C. (1998a). Computational models of colour constancy. In V. Walsh & J. Kulikowski (Eds.), *Perceptual constancy: Why things look as they do* (pp. 283–321). Cambridge, UK: Cambridge University Press.

Hurlbert, A. C. (1998b). Illusions and reality-checking on the small screen. *Perception, 27*, 633–636.

Ikeuchi, K., & B. K. P. Horn. (1981). Numerical shape from shading and occluding boundaries. *Artificial Intelligence*, *17*(1–3), 141–184.

Knill, D. C., & Kersten, D. (1991). Apparent surface curvature affects lightness perception. *Nature*, *351*, 228–230.

Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science*, *240*(4853), 740–749.

Logvinenko, A., & Menshikova, G. (1994). Trade-off between achromatic color and perceived illumination as revealed by the use of pseudoscopic inversion of apparent depth. *Perception*, *23*(9), 1007–1023.

Mach, E. (1959). *The analysis of sensation*s (S. Waterlow, Trans.). New York: Dover. (Original work published 1886)

Madison, C., Thompson, W., Kersten, D., Shirley, P., & Smits, B. (2001). Use of interreflection and shadow for surface contact. *Perception & Psychophysics*, *63*(2), 187–194.

Nayar, S. K., Ikeuchi, K., & Kanade, T. (1991). Shape from interreflections. *International Journal of Computer Vision*, *6*(3), 173–195.

Chapter 50
Color Assimilation

Frederick A. A. Kingdom

INTRODUCTION

Color assimilation, also known as the *Van Bezold spreading effect* (von Bezold, 1876; see one of von Bezold's original demonstrations in Evans, 1948, Plate XI, p. 193) is the phenomenon in which the perceived color of a region shifts *towards* that of its neighbor. Color assimilation is traditionally counter posed to color contrast, in which the color of a region appears to shift *away* from that of its neighbor. Figure III.50-1 shows an example of color contrast (left) and color assimilation (right). In the color-assimilation figure, the background regions appear tinted with the color of the overlaid grid. As a rule, however, the induction effects in such stimuli are weak. Stronger color assimilation effects can be observed in more complex stimulus arrangements, and three of these—the "special cases"—are discussed below. They are the chromatic version of White's effect (White, 1979), shown in Figure III.50-2, Monnier and Shevell's (2003, 2005) ring patterns, shown in Figure III.50-3, and neon color spreading (Anderson, 1997; Bressan, Mingolla, Spillmann, & Watanabe, 1997; van Tuijl, 1975), shown in Figure III.50-4. Finally, the watercolor illusion (Pinna, 2008; Pinna, Werner, & Spillmann, 2003), described in chapter 20 in this volume, may also be considered an example of color assimilation.

A NOTE ON TERMINOLOGY

The term "color" as used here refers to both the chromatic ("colored") and achromatic ("black-and-white") properties of a surface. The perceptual correlates of the terms chromatic and achromatic are respectively hue and lightness. Brightness, the perceptual correlate of luminance, or light intensity, is interchangeable with lightness in situations where there is no visible illumination component, and is used here wherever the literature uses the term. The examples of color assimilation described here are all in the chromatic domain, even though the causes of chromatic and achromatic color assimilation are closely related.

ASSIMILATION VERSUS CONTRAST

Although the definition of color assimilation is straightforward, the designation of a stimulus as one exhibiting assimilation is not straightforward, because a simple change in stimulus properties may cause a stimulus to switch from assimilation to contrast or vice versa. Following early work by Musatti (1953) and Kanizsa (1954), Helson and Rohles (1959) and Helson (1964) demonstrated quantitatively the change between contrast and assimilation in a simple stimulus comprising a series of lines on a grey background. They measured the difference in the lightness (perceived shade of grey) of the grey background depending on whether the overlaid lines were white or black. By varying the widths of both the lines as well as the intervening grey areas, they were able to map out the transitions from assimilation to contrast. They found that widening the lines caused a shift from assimilation to contrast, but the line width at which the transition occurred depended on the width of the grey regions in between. The important lesson from Helson's studies is that assimilation should not be identified with a particular stimulus or stimulus class but rather with a particular perceptual outcome.

CAUSES OF ASSIMILATION

Neural Blurring

The most widely entertained explanation of color assimilation is that it results from neural blurring. Visual stimuli are transduced by neurons that average the light signals that fall within their receptive fields. According to the neural blurring hypothesis, color assimilation occurs when the stimulus features are small compared to the size of the receptive-field subregions, with the result that the colors become averaged, or "spread out," within each subregion (Jameson, 1985). This explanation, however, needs careful consideration. The fact that our visual system possesses neurons that spatially average the colors that fall within their receptive-field subregions is not sufficient to explain color assimilation. Visual neurons are responsive to a wide range of spatial scales (and orientations); that is, they have large-, medium-, and small-sized receptive fields (see examples in Fig. III.50-3f–g). If the number and range of these neurons were sufficient to capture all the information in the image, constituting what is termed a "complete" basis set, and if the neurons operated as linear devices, then the totality of their responses would signal a veridical

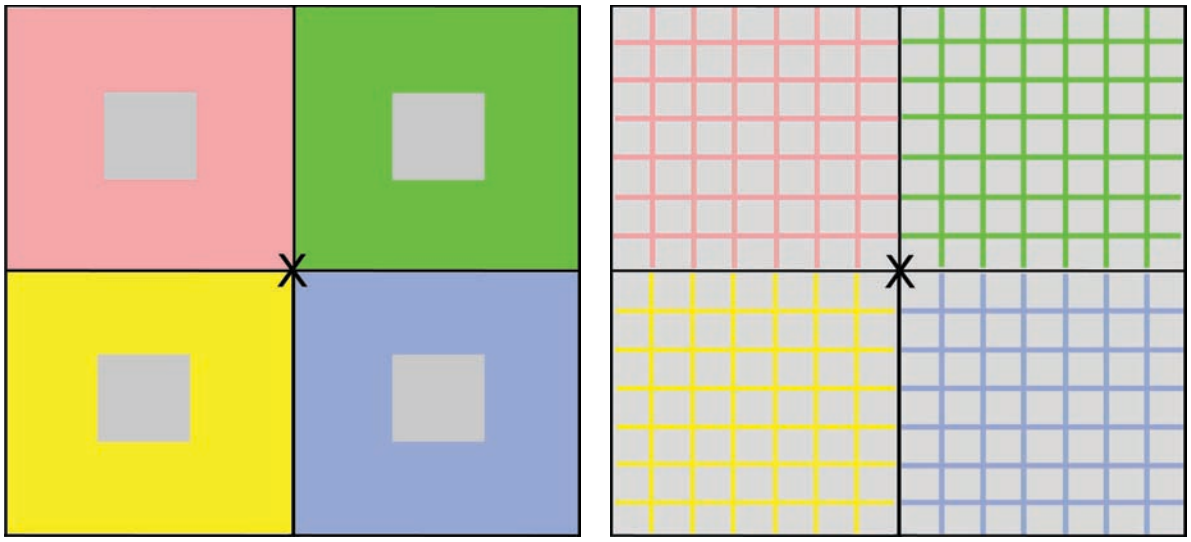


Figure III.50-1. Left: Simultaneous color contrast. The four grey squares appear tinted with desaturated colors complementary to those of their surrounds (clockwise from top right: cyan, pink, violet, and yellow). Right: Color assimilation. The grey backgrounds appear to be tinted with the colors of the overlay grids. Fixation on the central crosses might enhance the effects.

(i.e. accurate) representation of the visual input, and illusions such as color assimilation would not occur (Blakeslee & McCourt, 2004; Kingdom, 2011). Why? Because the response of a given scale of neuron serves to partially cancel out the effect of spatial averaging in the response of a larger-in-scale neuron, such that with a complete linear basis set there is no net spatial averaging of the input. The fact that we experience effects such as color assimilation therefore means one of three things: (a) the set of neurons that transduce the image is "incomplete"; in other words, some scales or orientations are missing; (b) the neurons' responses are subject to a nonlinear transformation; or (c) both (a) and (b). In what follows I examine how these possibilities can, in principle, lead to color assimilation.

First, how might an incomplete basis set of neurons lead to color assimilation? If the neurons sensitive to very small scales (i.e., those sensitive to very fine detail) are missing, color assimilation will occur. Without fine-scale neurons, there is nothing to counteract spatial averaging in the responses of the larger-scale neurons: let us call this the "missing fine-scale neurons" explanation of color assimilation. Interestingly, an analogous "missing coarse-scale neurons" explanation has been advanced for simultaneous brightness contrast (Blakeslee & McCourt, 2005; Kingdom, 2003; Shapiro & Lu, 2011), the achromatic version of simultaneous color contrast shown in Figure III.50-1. It is unclear whether the absence of fine-scale neurons is a significant factor in achromatic assimilation, since luminance-sensitive neurons found in the retina are of sufficient fine-scale sensitivity to capture the relevant information in the image, at least up to the limit set by the optical resolution of the eye. However, the absence of fine-scale neurons is likely to be a contributory factor in chromatic assimilation: chromatic vision is relatively insensitive to fine detail (reviewed in Shevell & Kingdom, 2008).

An important question arises from the missing fine-scale neurons explanation: Why do color assimilation stimuli not appear blurred? A likely reason is that the perception of blur and the perception of surface hue and brightness, are mediated by distinct neural processes. This idea was

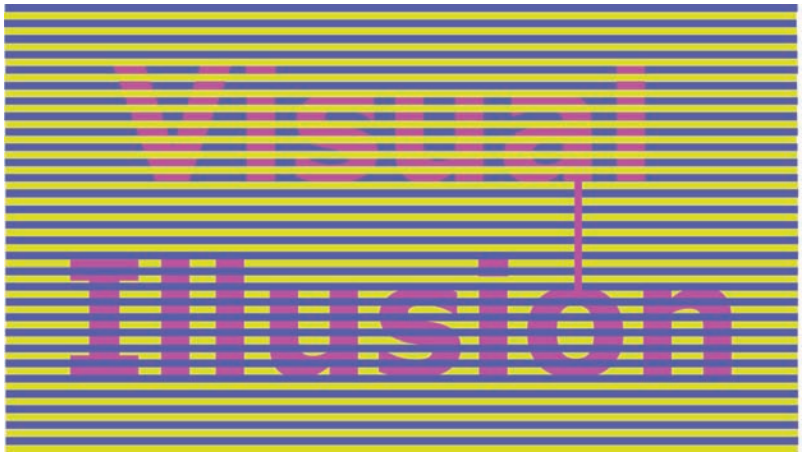


Figure III.50-2. Chromatic version of White's effect. The short horizontal bars that define the words "visual" and "illusion" have the same physical color (magenta) yet look remarkably different in hue due to the fact that they are interspersed by different phases of the background grating. (Based on Fig. 2 in Shevell & Kingdom, 2008.)

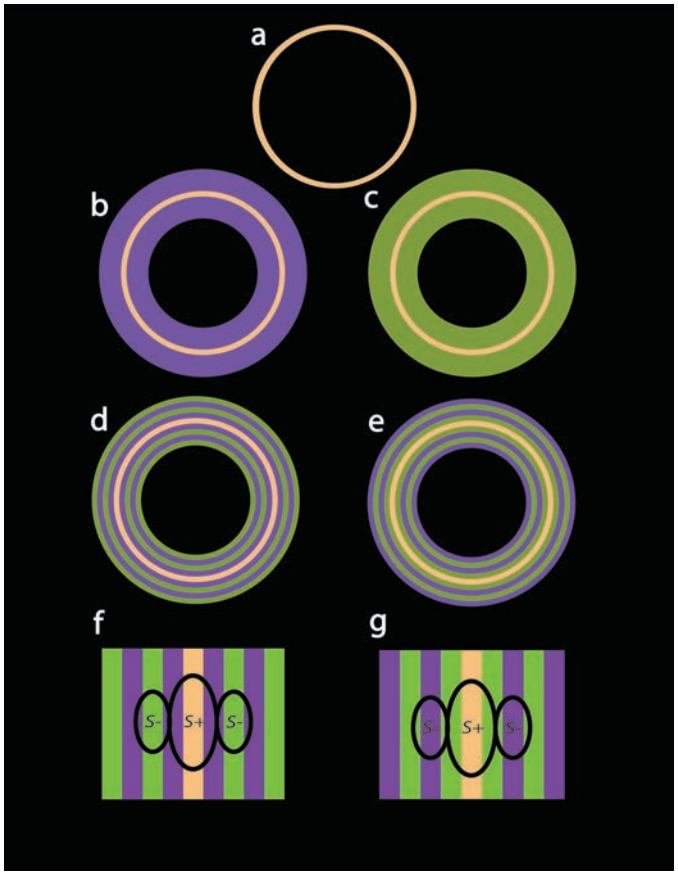


Figure III.50-3. Monnier and Shevell's (2004) ring patterns and proposed explanation for the illusory differences in hue. The ring shown in (a), when embedded in the circular lime-purple gratings (d) and (e) appears to be markedly different in hue depending on the phase of the grating it replaces (lime phase replaced on left, purple phase replaced on right). The effect is not simply due to assimilation from the neighboring bars, as shown in (b) and (c), which show a reduced effect; (f) and (g) show the explanation based on a center-surround receptive field with antagonistic S-cone inputs. (Based on Fig. 1 in Monnier & Shevell, 2004. Reproduced with permission of the authors.)

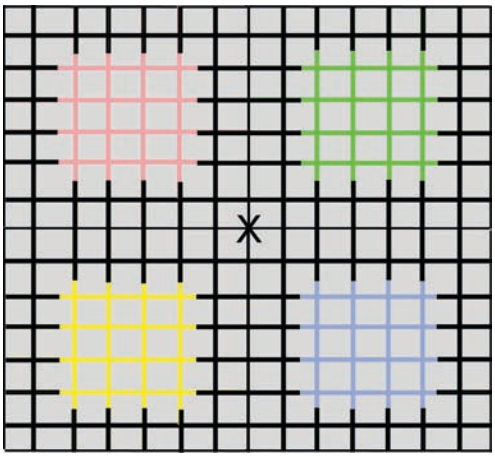


Figure III.50-4. Examples of neon color spreading.

implicitly suggested by Hurvich (1981), who in arguing against an *optical* spreading explanation for assimilation noted that "The individual elements of the pattern are seen as discreet elements (stripes or curves) despite the color blending. . . . there is a blending of hues and brightnesses (assimilation) but a maintenance of spatial pattern resolution" (p. 175). The separation of blur and surface-color perception is especially pertinent to chromatic assimilation. Typically, chromatic assimilation stimuli are not "isoluminant," in that they are not only defined by variations in chromaticity but also in luminance. Anecdotal evidence suggests that perceived blur may be signaled predominantly by luminance-sensitive neurons (see Wandell, 1995, Plate 7), in which case the perceived sharpness of the edges in the stimuli shown here may be signaled by the luminance system, whereas the color assimilation effects may result from neural averaging within the chromatic system.

Contrast Normalization

The second possible cause of color assimilation is that there is a nonlinear transformation applied to the early neural image, the likely candidate being contrast normalization. Contrast normalization is the canonical process whereby the visual system adjusts the relative gains, or response strengths, of neurons sensitive to different scales and orientations within a particular region of the image, in order to make the response strengths more equal. It has been suggested that the response-equalization property of contrast normalization serves to improve coding efficiency (Carandini & Heeger, 2012). Stimuli that exhibit color assimilation typically have a preponderence of fine-scale information. For example in Figure III.50-1, the most significant responses will be in the fine-scale neurons that detect the grid lines, while the coarse-scale neurons will respond relatively weakly to the regions in between. Contrast normalization might serve to shift the weighting of the signals from the fine- to the coarse-scale neurons, and because the latter neurons blend the colors of the contours with the grey color in between, the result is color assimilation. The idea that contrast normalization is a factor in color assimilation has been suggested by Blakeslee and McCourt (1999, 2004), Dakin and Bex (2003), and Robinson, Hammon, and de Sa (2007) in the context of the achromatic White's effect and its many variants, as discussed later.

Layer Decomposition

A third possible cause of color assimilation is layer decomposition. The visual system decomposes the image into different layers, and three of the most important layers are reflectance (surface albedo and color), illumination (shadows, shading, and highlights), and transparency (fog, acetate, milky water; see review in Kingdom, 2008). Anderson (1997, 2001) has argued that assimilation phenomena such as White's effect and neon color spreading result from the process of layer decomposition, or "scission". The possible role of layer decomposition in producing both of the above phenomena is the topic of the next section.

SPECIAL CASE 1: CHROMATIC WHITE'S EFFECT

A dramatic example of color assimilation is the chromatic version of White's effect, one version of which is shown in Figure III.50-2. The figure comprises a horizontally oriented blue-yellow grating with the words "Visual Illusion" defined by short, horizontal magenta stripes. The magenta stripes in "Visual" replace segments of the blue phase of the grating while those in "Illusion" replace segments of the yellow phase. In spite of the fact that the magenta color is identical in both words, the hues appear markedly different. Magenta is a mixture of red plus a little blue. The stripes defining "Visual" are mainly surrounded by the yellow phase of the grating (only the ends of the stripes in "Visual" abut the blue phase). When yellow is added to magenta, it tends to cancel the blue component leaving only red, i.e. what we perceive. On the other hand, the stripes defining "Illusion" are mainly surrounded by the blue phase of the grating, which augments the blue component in the magenta causing it to appear purple. Both illusory hues are hence in the direction of assimilation.

White's effect is best known by its achromatic version (White, 1979), in which short grey stripes appear lighter when placed on the black compared to the white phases of a square-wave grating. However, various forms of the chromatic version shown here were published earlier by Wright (1969, Plate 6, p. 86) and Munker (1970); White's effect is sometimes termed the Munker–White effect. Wright attributed the origin of the effect to an unpublished doctoral theses by Gindy (1963) and M. J. Melkman, and suggested it was caused by

> "scattered light in the eye, chromatic aberration of the optical system of the eye, simultaneous contrast, eye movements and local adaptation."

All three of the causes of assimilation discussed in the previous section might contribute to the chromatic White's effect in Figure III.50-2. First, the absence of very fine-scale chromatic-sensitive neurons will result in a net blending of the magenta with the stripes that intersperse them. Second, contrast normalization likely contributes to the illusion, as Blakeslee and McCourt (2004) have suggested for the achromatic White's effect. How so? The most responsive neurons in the stimulus are those whose receptive fields are matched in scale and orientation to the yellow and blue bars of the horizontal grating; these neurons will have relatively small, horizontally oriented receptive fields. Contrast normalization has the effect of reducing the weights of these neurons relative to those sensitive to the coarser scale and nonhorizontal orientations in the image. These latter neurons blend the colors of the letter bars with the interspersed grating bars, and because of their added weight, the blended colors are given prominence, resulting in the illusion.

Third, layer decomposition might be a contributory factor. Anderson (1997, 2001) has suggested that the process of scission underpins the achromatic version of White's effect. Translated to Figure III.50-2, the explanation is as follows. The junctions formed by the ends of the letter bars with the two grating phases, known as T-junctions, facilitate a process of perceptual scission, such that the letter stripes are treated as if part of a thin, transparent veil overlaying the grating bars. According to this account, the magenta color of "Visual" is perceptually "split" into its blue and red components, the blue being allocated to the grating stripes that run continuously underneath and the red to the veil. On the other hand, the magenta color in "Illusion" is split into yellow grating stripes and a purple veil. Since the letters of each word are defined by the veil that is generated by the process of scission, the words appear, respectively, red and purple.

The arguments for and against one or other, or some combination of these three explanations—or for that matter the many other explanations for White's effect and its chromatic version that have been put forward (e.g., Kingdom & Moulden, 1991; Moulden & Kingdom, 1989; Spehar, Clifford, & Agostini, 2002; Todorović, 1997; also see the brief review in Kingdom, 2011)—are beyond the remit of this chapter; the interested reader is referred to the cited papers.

SPECIAL CASE 2: MONNIER AND SHEVELL'S RING PATTERNS

A close relative of the chromatic White's effect is the ring pattern stimulus created by Monnier and Shevell (2003, 2004), shown in Figure III.50-3. A yellow-orange test ring (Fig. III.50-3a) is embedded in either the lime (Fig. III.50-3d) or purple (Fig. III.50-3e) phases of a circular lime-purple grating, resulting in dramatic differences in hue. The hue of the test ring shifts toward that of the adjacent ring's chromaticity. Color blurring is insufficient to explain the magnitude of the illusion; if one surrounds each test ring with a uniform field of the same chromaticity as the corresponding adjacent ring in the ring pattern (Figures III.50-3b and III.50-3c), the illusion is significantly reduced.

Monnier and Shevell (2003, 2004) suggest that the assimilative effects from the adjacent rings are augmented by contrast effects from the nonadjacent rings. The purple and lime colors stimulate primarily the short wavelength sensitive "S" cones of the human color vision pathway, with the purple eliciting positive and the lime negative S-cone responses. Monnier and Shevell propose that cortical neurons whose receptive fields have spatially antagonistic S-cone inputs (Conway, 2001) cause the shifts in color, as illustrated in Figures III.50-3f and III.50-3g. In Figure III.50-3f the response of the neuron is increased by +S-cone stimulation from the purple adjacent to the test ring and increased further by –S-cone stimulation from the nonadjacent lime ring, causing the test ring to shift dramatically toward purple. In Figure III.50-3g the combination produces instead a strong negative response, causing an opposite shift toward lime.

This explanation for the hue shifts in Monnier and Shevell's (2003, 2004) ring patterns is in keeping with the "missing fine-scale neurons" explanation for color assimilation described previously; if very fine-scale S-cone sensitive neurons were present, they would tend to counteract the shifts produced by the neurons whose receptive fields are shown in the figure.

SPECIAL CASE 3: NEON COLOR SPREADING

A particularly delicate form of color assimilation is the phenomenon known as neon color spreading (Anderson, 1997; Bressan et al., 1997; van Tuijl, 1975). The examples in Figure III.50-4 are simple extensions of the color assimilation pattern in Figure III.50-1b. The difference is that in Figure III.50-4 the ends of the colored grid lines abut with those from the black grid. The junctions between the two colours of grid lines are known as I-junctions. Because I-junctions occur whenever line contours are partially occluded by transparencies, they are believed to be important cues for perceptual transparency (Anderson, 1997). According to Anderson, I-junctions instantiate a process of perceptual scission, resulting in the decomposition of the image into two layers: a background layer consisting of line contours and a foreground layer consisting of a thin, colored, transparent overlay. As with the chromatic White's effect, however, neural averaging and contrast normalization might also contribute to the effect (e.g., see Robinson et al., 2007).

REFERENCES

Anderson, B. L. (1997). A theory of illusory lightness and transparency in monocular and binocular images. *Perception*, *26*, 419–453.

Anderson, B. L. (2001). Contrasting theories of White's illusion. *Perception*, *30*, 1499–1501.

Blakeslee, B., & McCourt, M. E. (1999). A multiscale spatial filtering account of the White effect, simultaneous brightness contrast and grating induction. *Vision Research*, *39*, 4361–4377.

Blakeslee, B., & McCourt, M. E. (2004). A unified theory of brightness contrast and assimilation incorporating oriented multiscale spatial filtering and contrast normalization. *Vision Research*, *44*, 2483–2503.

Blakeslee, B., & McCourt, M. E. (2005). A multiscale explanation of gradient induction and remote brightness induction effects: A reply to Logvinenko (2003). *Perception*, *34*, 793–802.

Bressan, P., Mingolla, E., Spillmann, L., & Watanabe, T. (1997). Neon color spreading: A review. *Perception*, *26*, 1353–1366.

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, *13*, 51–62.

Conway, B. R. (2001). Spatial structure of cone inputs to color cells in alert macaque primary visual cortex (V-1). *The Journal of Neuroscience*, *21*, 2768–2783.

Evans, R. M. (1948). *An introduction to color*. New York: Wiley.

Gindy, S. S. (1963) Techniques for subjective colour measurement and their application to colour contrast phenomena. Ph.D. Thesis, London University.

Helson, H. (1964). *Adaptation-level theory: An experimental and systematic approach to behavior*. New York: Harper & Row.

Helson, H., & Rohles, F. H. Jr. (1959). A quantitative study of the reversal of classical lightness-contrast. *The American Journal of Psychology*, *72*, 530–538.

Hurvich, L. M. (1981). *Color vision*. Sunderland, MA: Sinauer Associates.

Jameson, D. (1985). Opponent-colours theory in the light of physiological findings. In D. Ottoson & S. Zeki (Eds.), *Central and peripheral mechanisms of colour vision* (pp. 83–102). London: Macmillan.

Kingdom, F. A. A. (2003) Levels of brightness perception. In L. Harris & M. Jenkin (Eds.), *Levels of perception*. New York: Springer-Verlag.

Kingdom, F. A. A. (2008). Perceiving light versus material. *Vision Research*, *48*, 2090–2105.

Kingdom, F. A. A. (2011). Lightness, brightness and transparency: A quarter century of new ideas, captivating demonstrations and unrelenting controversy. *Vision Research*, *51*, 652–673.

Kingdom, F., & Moulden, B. (1989). Corner effect in induced hue: evidence for chromatic band-pass filters. *Spatial Vision*, *4*, 253–266.

Kingdom, F., & Moulden, B. (1991). White's effect and assimilation. *Vision Research*, *31*, 151–159.

Monnier, P., & Shevell, S. K. (2003). Large shifts in color appearance from patterned chromatic backgrounds. *Nature Neuroscience*, *6*, 801–802.

Monnier, P., & Shevell, S. K. (2004). Chromatic induction from S-cone patterns. *Vision Research*, *44*, 849–856.

Moulden, B., & Kingdom, F. (1989) White's effect: A dual mechanism. *Vision Research*, *29*, 1245–1256.

Munker, H. (1970). *Farbige gitter, abbildung auf der net- zhaut und übertragungstheoretische beschreibung der farbwahrnehmung* [Colored bars, the figure on the retina and description of transmission theory for color vision]. Munich: Ludwig-Maximilians-University of Munich.

Musatti, C. (1953). Colore e luce [Color and light]. *Archivo di Psicologia, Neurologia e Psichiatria*, *15*(5), 544–577.

Pinna, B. (2008). Watercolor illusion. *Scholarpedia*, *3*(1), 5352. http://www.scholarpedia.org/article/Watercolor_illusion

Pinna, B., Werner, J. S., & Spillmann, L. (2003). The watercolor effect: A new principle of grouping and figure–ground organization. *Vision Research*, *43*, 43–52.

Robinson, A. E., Hammon, P. S., & de Sa, V. R. (2007). Explaining brightness illusions using spatial filtering and local response normalization. *Vision Research*, *47*, 1631–1644.

Shapiro, A. G., & Lu, Z.-L. (2011). Relative brightness in natural images can be accounted for by removing blurry content, *Psychological Science*, *22*(11), 1452–1459.

Spehar, B., Clifford, C., & Agostini, T. (2002). Induction in variants of White's effect: Common or separate mechanisms? *Perception*, *31*, 189–196.

Todorović, D. (1997). Lightness and junctions. *Perception, 26*, 379–394.

van Tuijl, H. F. J. M. (1975). A new visual illusion: Neonlike color spreading and complementary color induction between subjective contours. *Acta Psychologica*, *39*, 441–445.

Von Bezold, W. (1876). *The theory of color in its relation to art and art-industry*. Boston: L. Prang.

Wandell, B. A. (1995). *Foundations of vision*. Sunderland, MA: Sinauer Associates.

White, M. (1979). A new effect on perceived lightness. *Perception*, *8*, 413–416.

Wright, W. D. (1969). *The measurement of color*. New York: Van Nostrand Reinhold.

Chapter 51
When Light Looks Like Paint

Frederick A. A. Kingdom

INTRODUCTION

> Suppose the sunlight falls on a small area of my black coat through a hole in the window shade of a room illuminated by another window: I see a gray spot on the coat and want to dust it off. (Hering, 1874/1964, p. 8)

This chapter describes misperceptions of light. Occasionally, nonuniform illumination such as a cast shadow, shading, or a spotlight appears to be paint or stain, or even a familiar object: in short, light sometimes appears to be "material." Artists have long been interested in the differences between material and illumination in their search for ways to depict shadows and shading, and vision scientists continue to explore how the brain distinguishes between these fundamental image dimensions, aided by carefully crafted laboratory stimuli (reviewed by Kingdom, 2008). This chapter, however, is concerned with situations where real light, rather than depictions of it in art or on display monitors, appears to be material. The illusions described below are not, however, all accidents of nature: some of them, such as the examples from shadow art, are human inventions, though they are no less genuine misperceptions for being so.

Why does light sometimes appear to be paint? The examples in this chapter allude to a common cause. Shadows, shading, and spotlights are constrained by physics to particular geometric and photometric forms that are familiar to us through experience. If these natural constraints are violated, our visual system switches to the only plausible alternative interpretation, namely that they are material. Helmholtz (1866/1962) famously coined the term "unconscious inference" to explain such misperceptions, capturing the idea that illusory visual percepts do not require conscious reflection. In modern-day parlance, the illusions in this chapter may be considered examples of Bayesian inference, the approach to understanding vision based on Bayes' theorem that incorporates the idea of prior knowledge via a formal probabilistic framework (e.g., Kersten, Mamassian, & Yuille, 2004). From the standpoint of Bayesian inference, our in-built knowledge about the conditions in which nonuniform illumination normally occurs leads us to assign a probability p to a stimulus that it is nonuniform illumination and by extension a probability $1 - p$ that it is material. Under certain conditions, $1 - p$ is sufficiently large that we incorrectly perceive the stimulus as material.

We rarely confuse shadows for paint in the natural visual world because natural scenes are replete with cues as to what is light and what is material. The examples reviewed here are the exceptions to the rule. However, from the standpoint of understanding how vision works, they are pivotal because they pinpoint the cues that vision exploits to distinguish between light and material.

SOME TERMINOLOGY

The forms of nonuniform illumination described in this chapter are shadows, shading, and spotlights. Shadows are caused by occlusion of nondiffuse illumination. Shading, sometimes referred to as "attached shadows," results from a change in the angle of a surface with respect to the direction of illumination: a surface that curves away from the direction of a light source becomes darker. A spotlight is the bright region of a surface caused by a localized beam of light.

The material properties described in this chapter are mainly surface reflectance and chromaticity. Reflectance is the percentage (or proportion) of light incident on an object that is reflected to the observer, and its perceptual correlate is lightness, or "perceived shade of grey." Thus a reflectance of 95% will tend to be perceived as white and a reflectance of 5% as black. Chromaticity refers to the "color" of a surface, strictly its spectral reflectance, or reflectance as a function of wavelength. The perceptual correlate of chromaticity is hue.

Two other important terms are luminance and contrast. Luminance is a measure of light intensity, and its perceptual correlate is brightness. Contrast, specifically luminance contrast, is the ratio of the luminance of a surface with its surround.

The distinction between lightness (perceived reflectance) and brightness (perceived luminance) is fundamental to this chapter. Thus a piece of white paper that is folded such that one half is less illuminated (i.e., in shade) may be perceived as uniform in lightness but non-uniform in brightness.

MACH CARD

The Mach card was first described in 1885 by the Austrian physicist Ernst Mach (Mach, 1885/1959), but is still studied (e.g., Bloj & Hurlbert, 2002). The Mach card is a convex (i.e., like a "<" when viewed from the left) folded card illuminated from one side, such that one side is brighter than the other (Fig. III.51-1). Normally both sides appear to be the same shade of grey, and, as Mach noted, observers "scarcely perceive" the brightness difference due to shading. However, prolonged viewing with one eye (to eliminate stereoscopic cues to three-dimensional shape) causes the shape of the card to reverse occasionally from convex to concave (i.e., like a ">" when viewed from the left). When this happens, the brightness difference becomes pronounced, and the dark side appears to be painted a darker shade of grey, instead of the same shade of grey but just less bright. The Mach card demonstrates how the interplay between lighting and perceived surface geometry can cause two surfaces with physically the same shade of grey to appear to have different shades of grey.

Our assumptions about the normal relationships between lighting and surface geometry explains why this happens. From the positions of the shadows in the image, we assume that lighting comes from the left and conclude that the brightness difference on either side of the card is attributable to lighting. In other words, we perceive the card as uniform in lightness (near-white) but with its right side in shade. However, when the card is perceived as concave, we would normally expect the right half to be more illuminated and hence brighter assuming light comes from the left. Since the right half is darker, the only reasonable conclusion is that it must be painted a darker shade of grey, and that is what we perceive.

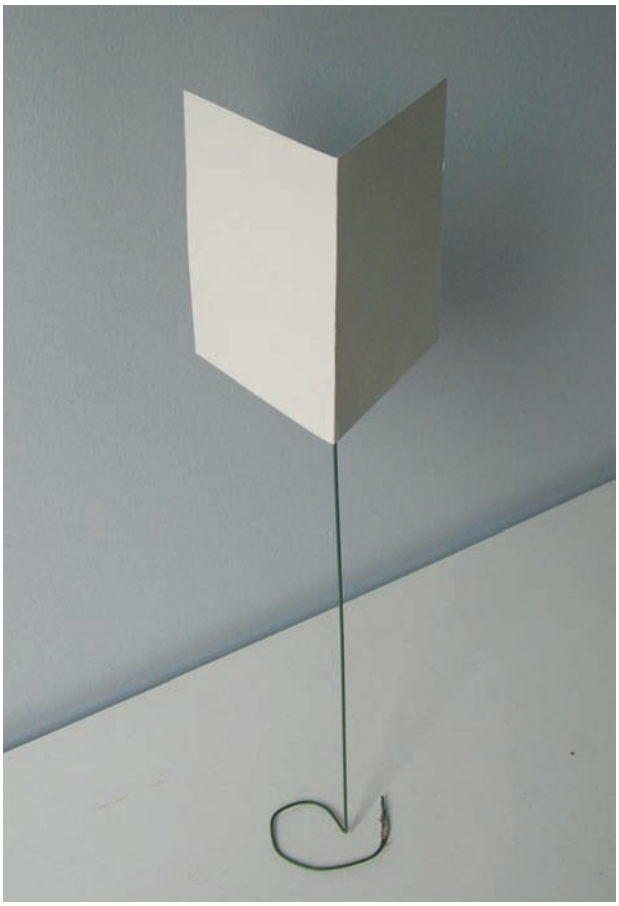


Figure III.51-1. Photograph of a Mach card. The white card is convex with respect to the camera and is illuminated from the left such that the right half is in shade. Prolonged fixation with one eye closed causes the card to perceptually reverse and appear concave. When this happens, readers may notice that the right half appears to be painted dark grey. (Photograph by Frederick Kingdom.)

Gilchrist (1977) noted that the outside corner of a building can sometimes appear flat when viewed monocularly, and when this happens the shadowed side appears to be of a darker material rather than of the same material but in shadow. In an interesting variant of this effect, Logvinenko and Menshikova (1994) used an optical instrument known as a pseudoscope to make a cone that cast a shadow appearing inverted in depth. With the inversion, the appearance of the shadow changed from shadow to paint.

Chromatic variants of the Mach card have been reported by Jakobsson, Bergström, Gustaffson, and Fedorovskaya (1997), and by Bloj, Kersten, and Hurlbert (1999). Bloj et al. used a Mach card shape to test whether interreflected rather than projected colored light may appear to be paint (see chapter 49 in this volume). They painted the inside of a Mach card magenta on one side and white on the other. Some of the light from the magenta half was reflected onto the white half, and as a result observers perceived a pinkish glow on the white side. However, when the perceived shape of the card was inverted with a pseudoscope, so that the card appeared convex, the white half of the card took on a deep magenta color. The shift in color percept accompanying the change in perceived shape can be attributed to the visual system's prior knowledge of physical interreflection. The reflected light from the magenta half that fell on the white half tended to be discounted in the normal view because the geometry was consistent with interreflection between surfaces. However, in the inverted-shape, convex view, the arrangement was not consistent with interreflection and the light reflected from the magenta half was attributed to the paint on the surface of the white half, which as a result appeared magenta.

HERING'S SHADOW ILLUSION

Another influential phenomenon that emerged from 19th-century vision science is Hering's (1874/1964) shadow illusion. If one draws a thick line along the penumbra of a cast shadow, the shadow's appearance changes from dark illumination to a different shade of grey. The illusion has been used to great effect in street art, as in Figure III.51-2 an example from the New York artist and photographer Michael Neff.

Why does this illusion happen? The likelihood that the edge of a cast shadow would normally coincide with an identically shaped line contour of material origin is extremely small, so the visual system infers that the dark region must be part of the same object as the line contour and hence be material. The chalk line also covers the shadow's "penumbra," or blurred edge, which is an important cue to a shadow (Kingdom, 2008). A number



Figure III.51-2. Hering's shadow illusion in art: a chalk outline around a shadow makes the shadow appear like chalk or paint. One of many examples by New York artist/photographer Michael Neff. (Reproduced with permission.)

of variants of Hering's shadow illusion have been used to study the precise conditions that give rise to this illusory effect (e.g., Cavanagh & Kennedy, 2000; Cavanagh & Leclerc, 1989; Kennedy & Bai, 2000).

GELB EFFECT

In the introductory quote to this chapter Hering notes how small spotlights sometimes appear to be paint or stain. Probably the best-known example of this illusion is the Gelb effect, popularized in recent times by Alan Gilchrist in numerous live-audience demonstrations (Gelb, 1929; a translation of Gelb's own description of the effect can be found in Gilchrist, 2006, p. 62). An example of the Gelb effect created in the author's laboratory is shown in Figure III.51-3. Figure III.51-3a shows what appears to be a light grey circular disk attached to a stand. The disk is in fact black but illuminated by a narrow beam of light from a projector on the ground below (not shown); the unoccluded portion of the beam can be seen falling on the air-conditioner unit above the window. In Figure III.51-3b the projector has been switched off and the true color of the disk becomes apparent. The effect is very compelling—even when one realizes that the light shade of grey of the disk is an illusion caused by the light from the projector, it is hard to shake the impression that it is light grey. The critical physical property underpinning the Gelb effect is that none of the light from the projector falls on the background against which the disk is viewed, so a discontinuity in illumination occurs precisely at the border of the disk with its background. In other words, a change in illumination (from the light from the projector beam on the disk to the light from the room/window lighting) is perfectly correlated with the change in reflectance (from the color of the disk to the color of the background). Normally we are able to identify an illumination change from various cues (such as from the crossing of a shadow and a reflectance change know as an X-junction) and are thus able to "discount" the effects of a projector beam and correctly identify a stimulus as black. However, in the absence of the normal visual cues for decoupling illumination edges from reflectance edges, the visual system reasonably concludes that the disk must be light grey because its luminance contrast with its background is so high. If one places a white card behind the illuminated disk, the disk immediately appears black, reinforcing the idea that the lightness of the disk is caused by contrast.

Probably the most common example of the Gelb effect in the natural visual world is the moon seen in daylight. As noted by Wallach (1976) and discussed recently by Bressan (2005, 2007), the moon in daylight appears white even though its surface is dark grey. In daylight the moon still reflects more of the sun's light than does the surrounding sky, causing a large contrast between moon and surround. At night the contrast becomes so high that the moon appears self-luminous, a "super-Gelb" effect.

SHADOW ART ILLUSION

The Azerberjan artist Rashad Alakbarov is well known for creating realistic figures of objects from shadows. In one piece, shown in Figure III.51-4, the suspended plastic bottles have been meticulously arranged such that their shadows form a cartoon-like face on the wall behind. What is striking is that the face appears to be chalk-drawn or spray-painted rather than composed of shadows. Again, prior knowledge of the constraints in which shadows normally occur is the explanation. The likelihood that a real face under normal lighting conditions would ever produce the pattern of shadows on the wall is negligible: shadows are formed by occlusion of light, so shadows of faces are invariably silhouettes (i.e., devoid of internal structure). Thus our visual system concludes that the face in Figure III.51-4 is not constructed from shadows but from chalk or paint.

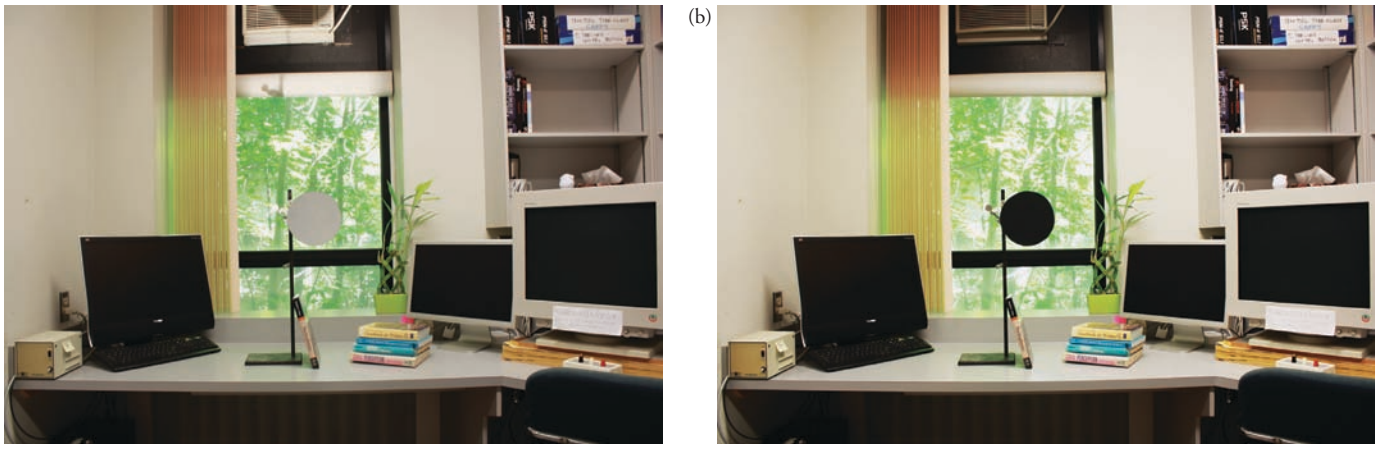


Figure III.51-3. The Gelb effect. (a) The circular disk attached to the stand in the middle of the photograph appears light grey. In fact, it is a black disk illuminated by a beam of light from a projector placed on the floor below (not shown). The bottom half of the shadow of the disk cast by the projector can be seen on the air-conditioner unit above the window. (b) The real shade of the disk is revealed when the lamp is switched off. (Photos by Marouane Ouhnana.)



Figure III.51-4. Shadow art by Rashad Alakbarov.

CHROMATIC SHADOW ILLUSIONS

The tower block in Figure III.51-5b appears to be painted orange, with the painters running out of paint about two-thirds of the way down. In fact the orange color is sunset illumination on a grey surface. The figure demonstrates how a vivid-colored illumination can be misperceived as paint. In Figure III.51-5c a wall that is painted light blue appears to be painted blue-green. In fact the greenish tint is due to green-colored illumination from light outside that is filtered through green foliage. Figure III.51-5d shows a garden exhibit in which the ground is covered by glass beads. The strongly colored circular patches on the ground appear to be stained



Figure III.51-5. (a) Grass-pavement border with tree shadow. (b) The orange color of the tower is sunset illumination on a grey surface. (Photo by Brian Micklethwait; reproduced with permission.) (c) Gaudy colors? The photo from the author's house shows a greenish-blue wall on the right, yet the wall is painted the same color blue as the narrow vertical wall that it abuts. The greenish tint comes from the light from outside, which is filtered through green foliage. (d) Garden exhibit showing what appears to be circles of colored glass yet are colored shadows, as revealed in (e). The exhibit is named "Between Earth and Sky" and was created by Juliette Patterson and Michel Langlois. (Photos of exhibit reproduced with permission.)



Figure III.51-6. A herd of camels photographed from above, such that only the shadows are easily discernible. (Photo by George Steinmetz; reproduced with permission.)

glass. Yet these are colored shadows cast from semitransparent disks suspended from above, as revealed in the photograph in Figure III.51-5e. Again, color that is light is mistaken for color that is stain.

Why then does colored light sometimes appear to be paint or stain? The reason is that our visual system has an inbuilt bias toward assuming that chromatic variations are material in origin, that is, arise from paint, stain, or pigment (Kingdom, 2003). In the natural visual world most chromatic variations are material, whereas luminance, or brightness variations, can be either material or illumination (Olmos & Kingdom, 2004; Rubin & Richards, 1982). Consider Figure III.51-5a, which shows the shadow of a tree falling across a border between grass and pavement. The grass–pavement border is primarily a change in three features: luminance (dark to light), color (green to grey), and texture (coarse to fine). The shadow, on the other hand, is primarily a change in luminance—there is no accompanying change in texture, and the color changes are usually small.[1] Thus while the color changes are primarily material in origin (grass to pavement), the luminance changes in the figure are both material (grass to pavement) and light (the shadow). Moreover, when the luminance change is material, it tends to coincide with color and textural changes, whereas when the luminance change is light, it tends to occur on its own. This simple figure illustrates an important constraint in the physical relations between color and luminance in the natural visual world: color changes, and those luminance changes that co-occur with them, tend to be material in origin, whereas luminance changes on their own tend to be caused by light. I term this the "color is material" assumption. Thus even when the colored region is light, such as sometimes occurs with shadows or shading, we are biased toward it appearing material.

SHADOWS THAT APPEAR AS OBJECTS

In extreme circumstances, shadows can appear to be familiar objects, as in Figure III.51-6. This photograph by George Steinmetz is of a herd of camels taken from high above, such that the camels themselves are hardly visible because of the angle they subtend to the camera. The shadows look real because there are no visible objects that appear to be casting them.

NOTE

1. On bright sunny days deep shadows tend to be bluish. This is because the shadow receives only blue skylight whereas the unshadowed part receives a mixture of yellow sunlight and blue skylight.

REFERENCES

Bloj, M. G., & Hurlbert, A. C. (2002). An empirical study of the traditional Mach card effect. *Perception, 31*, 233–246.

Bloj, M. G., Kersten, D., & Hurlbert, A. C. (1999). Perception of three-dimensional shape influences colour perception through mutual illumination. *Nature*, *402*, 877–879.

Bressan, P. (2005). The dark shade of the moon. *Clinical and Experimental Ophthalmology*, *33*, 574.

Bressan, P. (2007). *Il colore della luna: Come vediamo e perché* [The color of the moon: How we see and why]. Rome: Laterza.

Cavanagh, P., & Kennedy, J. M. (2000). Close encounters: Details veto depth from shadows. *Science*, *287*, 2423–2424.

Cavanagh, P., & Leclerc, Y. (1989). Shape from shadows. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 3–27.

Gelb, A. (1929). Die "farbenkonstanz" der sehdinge [The color constancy of visual objects]. In W. A. von Bethe (Ed.), *Handbuch*

der normalen und pathologischen physiologie (Vol. 12, pp. 594–678). Berlin: Springer.

Gilchrist, A. (2006). *Seeing black and white*. Oxford: Oxford University Press.

Gilchrist, A. L. (1977). Perceived lightness depends on perceived spatial arrangement. *Science*, *195*, 185–187.

Helmholtz, H. von. (1962). *Treatise on physiological optics*. Vol. II (J. P. L. Southall, Trans.). New York: Dover. (Original work published 1866)

Hering, E. (1964). *Outlines of a theory of light sense*. Cambridge MA: Harvard University Press. (Original work published 1874)

Jakobsson, T., Bergström, S. S., Gustaffson, K.-A., & Fedorovskaya, E. (1997). Ambiguities in colour constancy and shape-from-shading. *Perception*, *26*, 531–541.

Kennedy J. M., & Bai, J. (2000). Cavanagh and Leclerc shape-from-shadow pictures: Do line versions fail because of the polarity of the regions or the contour? *Perception*, *29*, 399–408.

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology*, *55*, 271–304.

Kingdom, F. A. A. (2003). Colour brings relief to human vision. *Nature Neuroscience*, *6*, 641–644.

Kingdom, F. A. A. (2008). Perceiving light versus material. *Vision Research*, *48*, 2090–2105.

Olmos, A., & Kingdom, F. A. A. (2004). A biologically-inspired algorithm for separating reflectance and shading images. *Perception*, *33*, 1463–1473.

Rubin, J. M., & Richards, W. A. (1982). Color vision and image intensities: When are changes material? *Biological Cybernetics*, *45*, 215–226.

Wallach, H. (1976). *Hans Wallach on perception*. New York: Quadrangle/New York Times Book Co.

Chapter 52
The Scintillating Grid

Bernd Lingelbach

DISCOVERY OF THE SCINTILLATING GRID EFFECT

The first presentation of the scintillating grid was made in August 1995 at the European Conference on Visual Perception (ECVP) in Tübingen, where it immediately attracted a great deal of attention (Schrauf, Lingelbach, Lingelbach, & Wist, 1995). At the time, the Hermann grid seemed to have been completely explained. It was actually the logo of the conference and adorned the title page of the volume of conference abstracts. Spillmann (1994) had just published his review of the Hermann grid, and no one was expecting anything new. But the minor modification of having a grey instead of a white grid with white dots at the intersections produced a strikingly new and powerful illusion. Instead of the white dots, dark black dots are seen to "blink" and appear even darker in intensity than the black printer's ink of the square (Fig. III.52-1, left). They then disappear again to reappear immediately somewhere else. The black dots appear only in the peripheral field of vision: at the central point of focus, only white dots are visible. By varying the observation distance, it is easy to discover that there is an individual optimal distance at which the effect is strongest. It also works when viewed with only one eye, but in this case there are far fewer flickering black dots. The illusion can be reversed as well, still using grey bars but with white squares producing black dots (Fig. III.52-1, right). A strong scintillating effect can also be achieved with suitable color combinations.

Not long after the conference, a full report was published in *Vision Research* (Schrauf, Lingelbach, & Wist, 1997). For a topical reason, in the year 2000 the pattern became known to a wider audience when it was sent winging round the world by e-mail from computer to computer as an ironic "set of instructions" for counting the votes after the presidential elections, which had hitherto proved elusive: "Florida Election Recount. Count and total the black dots for Al Gore and the white dots for George Bush. Recount to confirm."

WHAT LED TO THE SCINTILLATING GRID EFFECT?

The Hermann grid effect first attracted the attention of perception researchers when they thought they could explain it. The concept of receptive fields appeared to offer a perfect explanation of the blobs (Baumgartner, 1960). But arguments had surfaced early on that countered this explanation. For example, there were the dark diagonals that Motokawa (1950) had already spotted 10 years earlier in his inclined version of the Hermann grid. They could be explained much better by the concept of spatial frequencies. One particular example (around 1980) did not fit with Baumgartner's explanation: in an array of black circles, the white areas in between should appear grey according to the receptive field theory. But in fact the transitions from circle to circle are perceived as greyer than the white paper, just as low-pass filtering would predict. At a suitable distance, a blurred dark grid pattern can even be seen, where the lines pass through the centres of the circles and appear darker than the black of the circles. With the Hermann grid, sometimes dark diagonals can be seen that are actually there in the low-pass filtered version of the grid. These dark diagonals appear to explain the grey spots at least as well as assumptions about appropriate receptive fields in

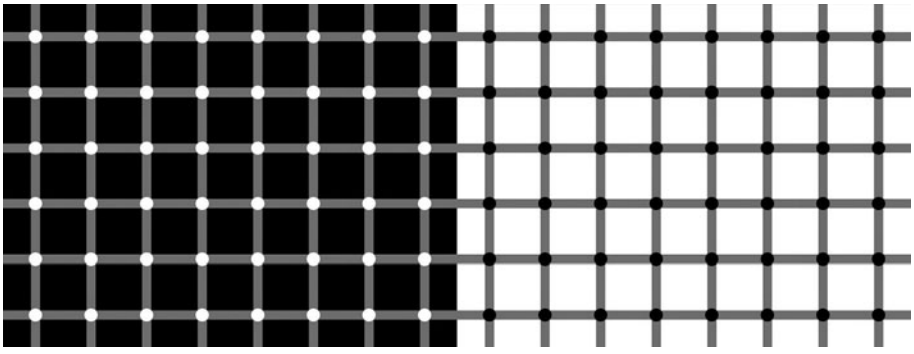


Figure III.52-1. Scintillating grid, positive and negative. Instead of the white dots (on the left) dark black dots appear to blink at the periphery. In the negative grid (on the right) precisely the opposite happens. In place of the black dots, bright white dots appear. (Image: Lingelbach B & E, 1995.)

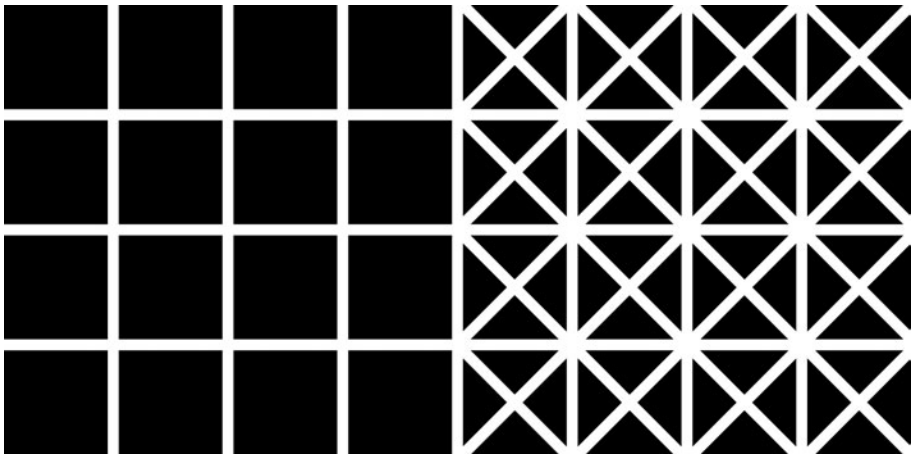


Figure III.52-2. Left: Hermann grid. Right: Grid with additional diagonal lines drawn in. The Hermann grid illusion has now disappeared (Image: Lingelbach, 1985).

the retina: at each white intersection the dark diagonals cross; the crossing should thus be perceived as darker.

It was just a short step to take out the dark diagonals (Fig. III.52-2). Indeed, the grey spots did then disappear. Figure III.52-2 was actually shown at ECVP 1985, but at that time it was not taken seriously as evidence against the then-prevailing explanation of the Hermann grid (Lingelbach, Bock, Hatzky, & Reisinger, 1985). Only much later was the pattern taken up again (Schiller & Carver, 2005).

The explanation accepted at that time was based on the understanding of concentric receptive fields that are typical for the retina. If these alone were responsible for the Hermann grid illusion, then the illusion should not be affected by orientation. However, it can be observed, on the contrary, that with a Hermann grid at 45° the brightness of the illusion becomes weaker. Furthermore, there is another clear effect: black diagonal lines can be perceived. These were first observed by Prandtl (1927) on a Hermann grid and described as "spider's web threads." The dependence of the illusion on orientation speaks in favor of the influence of orientation-selective neurons with longitudinal receptive fields and—as these are only to be found in the cortical plane—against a purely retinal source for the illusion. These spider's web lines are similar to the dark lines in Motokawa's (1950) diamond-shaped variant of the Hermann grid. The shortcomings, or near impossibility, of an explanation based solely on retinal receptive field properties is shown very convincingly in Figure III.52-2, right. The additional white diagonal lines inserted here should lead to a strengthening of the illusion of black diagonal lines, due to less lateral inhibition. In fact they lead to their disappearance, or at least to their becoming extremely weak. A more detailed and excellent investigation on the Hermann grid and diagonal lines is given by De Lafuente and Ruiz (2004).

Fourier Analysis

The particular significance of the appearance of diagonals argues in turn for an explanation at the higher cortical plane of the visual system where orientation-specific processing of visual information takes place. Specifically, the diagonal effect satisfies the two-dimensional "Fourier model" of visual data processing that dates back to Campbell and Robson (1968). According to this, every two-dimensional pattern can be generated from a sum of suitable sine patterns (Fourier synthesis). Conversely, the pattern can be broken down into its components from sine patterns (Fourier analysis). Some evidence at the time suggested that in the visual system something like a Fourier analysis takes place. This too can easily be demonstrated on the Hermann grid (Fig. III.52-2, left). In the Fourier analysis, grids rotated through 45° clearly exhibit dark lines running through the intersections of the white bars. Perhaps it is these dark lines that are sometimes perceived as diagonal spider's web lines in the Hermann grid. These illusory lines can be seen most clearly in the horizontal–vertical directions (i.e., rotated at 45° to the Hermann grid). As is well known, orientation-specific neurons are more prevalent in these two directions than in the diagonals. The perception of horizontal and vertical structures is accordingly also more sensitive.

The Scintillating Grid

It thus seemed appropriate to examine the Hermann grid more closely using Fourier analysis. In the summer of 1994, Elke and Bernd Lingelbach experimented with a low-pass filtered version of the Hermann grid and observed the strange flashing of black dots (Fig. III.52-3). J. S. Bergen had already spotted this effect on a similar pattern in 1985 and coined the term "scintillating grid." However, the scintillation can be seen much more clearly in Figure III.52-3 than in Bergen's original.

Elke Lingelbach's crucial contribution was now to reduce the pattern in Figure III.52-3 to its key components: the black squares, the grey grid, and the white dots. This new pattern was the scintillating grid that was presented for the first time at ECVP in Tübingen in August 1995 and that has remained practically unchanged to this day.

Three-Dimensional Variant

In Figure III.52-4 the dots are horizontally slightly off center. However, the effect still persists. This shows that the illusion does not depend on the precise location of the dots. However, if the picture is viewed as a kind of "magic-eye picture" then the white dots appear to float in front of and behind the plane of the image. The illusion is still there despite it being considerably weaker. Now and again white dots appear to be black.

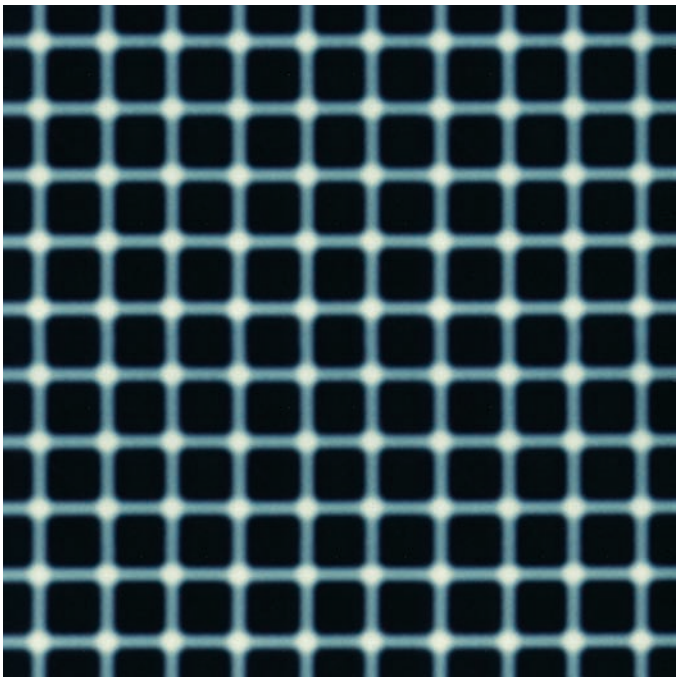


Figure III.52-3. Low-pass filtered Hermann grid. The light dots appear to "blink" black. A similar although much weaker effect was observed by J. R. Bergen in 1985 and named "the scintillation grid." (Image: Lingelbach B & E, 1995.)

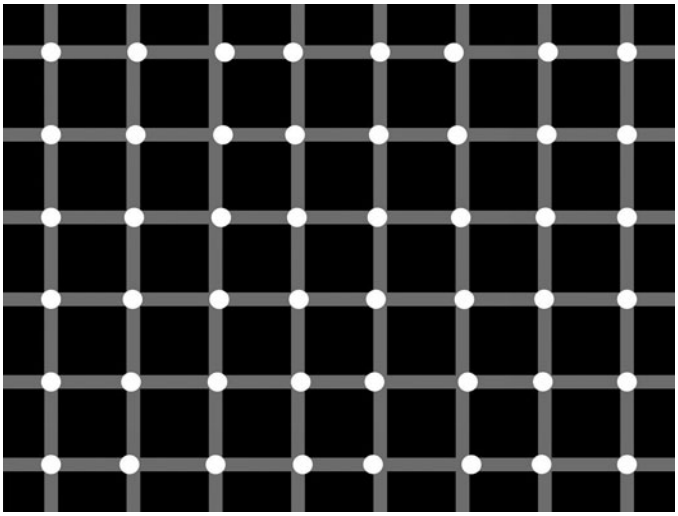


Figure III.52-4. A 3D-variant of the scintillating grid. The white dots are slightly displaced to the right and left. If the figure is now viewed as a "magic-eye image," the points appear to float in front of and behind the plane of the Hermann grid. Despite this the scintillating effect can still be seen. (Image: Lingelbach B, 2000.)

ATTEMPTED EXPLANATIONS

In the intervening years a deluge of literature on the scintillation effect appeared. However, no satisfactory explanation has so far emerged. The illusion is certainly more than just a contrast phenomenon, as the effect is much too strong and fluctuating. The illusion is also not necessarily tied to the regularity of the grid structure. It can occur also, although it is considerably weaker, when the intersecting elements of the grid are isolated and irregularly distributed, and it remains even when the lines do not cross (Heidler, Lingelbach, Ehrenstein, & Lingelbach, 2000). This calls for an explanation that is not exclusively based on a global Fourier analysis. At least local characteristics of the image analysis (possibly Gabor functions or wavelets) must be included too (Heidler, 1999).

The global Fourier analysis even breaks down with the regular scintillating grid in Figure III.52-1. Here, as with the Hermann grid, it is possible to calculate the Fourier components mathematically. While with the Hermann grid dark diagonals are actually present in the Fourier spectrum and combine to form dark areas at the intersections, this is absolutely not the case with the scintillating grid. We carried out calculations for a Hermann grid with black squares, grey bars, and variable brightness of the intersections. When the intersections have the same luminance as the bars, then indeed the grey Hermann spots do appear at the crossings in the low frequencies in accordance with the calculation. As the luminance of the intersections increases, the calculated Hermann spots should become weaker according to the Fourier components. In the case of white dots at the crossings, the Fourier components result in white diagonals, the exact opposite of what is perceived. If the visual system is really carrying out a Fourier analysis, then the maximum effect would be seen with the original Hermann grid. But this is not the case.

A paradoxical effect occurs, however, if the white dots are replaced by red dots. Astonishingly enough, red diagonals can be perceived (Fig. III.52-5), as predicted by the Fourier spectrum of the pattern.

A NEW GRID VARIANT

Investigation of many variations of the scintillation grid led to a further pattern, which was presented for the first time at ECVP 2002 in Glasgow (Lingelbach, Ehrenstein, & Lingelbach, 2002). The basic pattern is rotated through 45°, and on three horizontal and three vertical stripes the

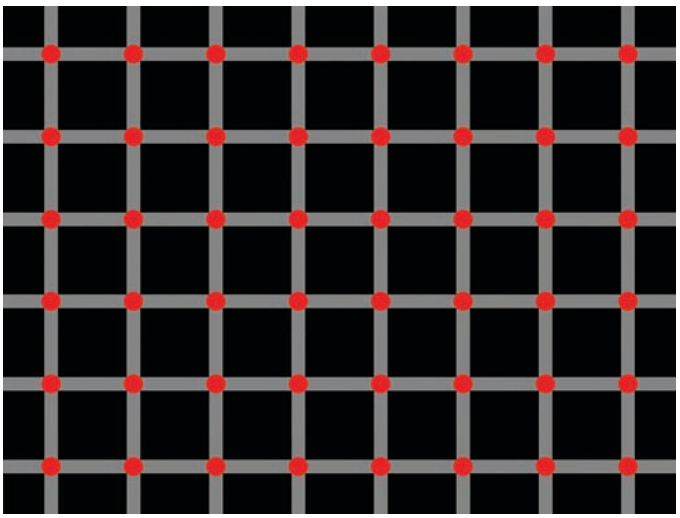


Figure III.52-5. The scintillating grid with red dots. The scintillating effect has nearly vanished. But the red dots seem to induce red diagonals. (Image: Lingelbach B, 2000.)

dots are simply omitted (Fig. III.52-6). The area where the dots are missing appears lighter than the objectively similar but apparently darker background of the square where the dots remain. Moreover, in the wide, lighter stripes three horizontal or vertical dark lines can clearly be seen. Their appearance is much more pronounced than in Motokawa's (1950) pattern. In a large poster the effect is so powerful that observers first believe that the dark lines are not just apparent but are actually there. On further observation an additional unique phenomenon occurs. The virtual lines start to shiver and flicker and seem to move about a bit. This additional dynamic quality of movement is reminiscent of an image from the "Enigma" series by Isia Leviant (1996), in which colored rings on a radial black-and-white background appeared after a certain time to start rotating "as if carried by a flow of material." How the dynamic effect of both the pulsating "flashes" and the flowing lines is created is still completely open to interpretation. What is certain is that eye movements play an important part here. But this alone, as with the explanation of the Hermann grid, is hardly enough because, above all other reasons, the trembling effect as well as the clarity of the perceived stripes critically depend on the particular color combinations of the pattern.

At ECVP 2004 in Budapest, Janos Geier et al. (2004) presented a simple yet very effective variant of the Hermann grid. The straight stripes are just slightly bent in the shape of a sine wave, and this has the effect of making the faint light grey spots of the Hermann illusion almost completely disappear. In addition, the scintillating grid effect disappears if the lines are bent.

This astonishing phenomenon shows that work on the Hermann grid and its variants will certainly continue. Suffice it to say here that the perceived lines in Figure III.52-6 also disappear when the grey lines are of similar sinusoidal shape to those in the pattern from Geier et al. (2008)

It remains to be seen whether the Geier pattern heralds a new (post-Geier) era. Whatever happens, textbook explanations of the Hermann grid will in the future need to be formulated more circumspectly.

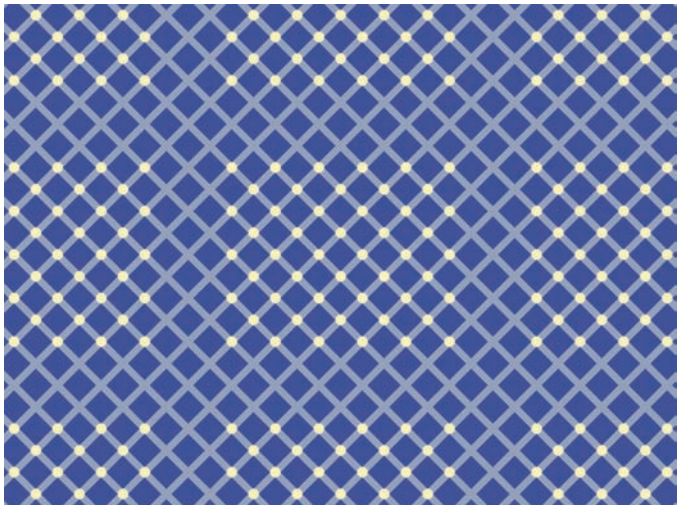


Figure III.52-6. Diagonal-orthogonal grid combination: In this pattern three further perception effects can be observed: (a) the dark background appears to be considerably lighter in the striped areas where the white dots have been left out; (b) within these wide lighter stripes, three dark horizontal or three dark vertical lines, respectively, can be seen; (c) after some time these illusory lines start to flicker, creating the impression of moving back and forth along the lines. (Image: Lingelbach B, 2002.)

ACKNOWLEDGMENTS

Translation from the original German by John Saniter with editorial help from Nick Wilson.

REFERENCES

Baumgartner, G. (1960). Indirekte Größenbestimmung der rezeptiven Felder der Retina beim Menschen mittels der Hermannschen Gittertäuschung. *Pflügers Archiv für die gesamte Physiologie*, *272*, 21–22.

Bergen, J. R. (1985). Hermann's grid: New and improved [Abstract]. *Investigative Ophthalmology and Visual Science*, *26*(Suppl.), 280.

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology*, *197*, 551–566.

De Lafuente, V., & Octavio Ruiz O. (2004). The orientation dependence of the Hermann grid illusion. *Experimental Brain Research*, *154*, 255–260.

Geier J., Sera L. & Bernath L. (2004): Stopping the Hermann grid illusion by simple sine Distortion. [Abstract]. *Perception*, *37*(Suppl.), 53.

Geier J., Bernath L., Hudak, M. & Sera L. (2008): Straightness as the main factor of the Hermann grid illusion. *Perception*, *37*, 651–665.

Heidler, A. (1999). *Fourieranalyse als Wahrnehmungsmodell diskutiert anhand der Scintillation-Grid-Täuschung*. Diplomarbeit, FB Augenoptik, FH Aalen.

Heidler, A., Lingelbach, E., Ehrenstein, W. H., & Lingelbach, B. (2000). Szintillationseffekt ohne Gitterstruktur. In H. H. Bülthoff (Ed.), *Beiträge zur 3. Tübinger Wahrnehmungskonferenz* (pp. 66–67). Kirchentellinsfurt, Germany: Knirsch.

Leviant, I. (1996). Does "brain-power" make Enigma spin? *Proceedings of the Royal Society B*, *263*, 997–1001.

Lingelbach, B., Bock, B., Hatzky, B., & Reisinger, E. (1985). The Hermann grid illusion—Retinal or cortical? [Abstract]. *Perception*, *14*, A7.

Lingelbach, B., Ehrenstein, W. H., & Lingelbach, E. (2002). A new Hermann grid variant: Scintillating dots, spurious lines, neon-colour spreading, vacillating and pulsating motion [Abstract]. *Perception*, *31*(Suppl.), 162.

Motokawa, K. (1950). Field of retinal induction and optical illusion. *Journal of Neurophysiology*, *18*, 413–426.

Prandtl, A. (1927). Über gleichsinnige Induktion und die Lichtverteilung in gitterartigen Mustern. *Zeitschrift für Sinnesphysiologie*, *58*, 263–307.

Schiller, P., & Carvey, C. E. (2005). The Hermann grid illusion revisited. *Perception*, *34*, 1375–1397.

Schrauf, M., Lingelbach, B., Lingelbach, E., & Wist, E. R. (1995). The Hermann grid and the scintillation effect. *Perception*, *24*(Suppl.), 88–89.

Schrauf, M., Lingelbach, B., & Wist, E. R. (1997). The scintillating grid illusion. *Vision Research*, *37*, 1033–1038.

Spillmann, L. (1994). The Hermann grid illusion: A tool for studying human perceptive field organization. *Perception*, *23*, 691–708.

Chapter 53
Second-Order Mach Bands, Chevreul, and Craik-O'Brien-Cornsweet Illusions

Zhong-Lin Lu and George Sperling

INTRODUCTION

When two plateaus of constant luminance are joined by a linear luminance ramp (Fig. III.53-1a), illusory bands are perceived at the junctions—an induced dark band is perceived at the bottom of the ramp and a bright band near the top of the ramp (Fig. III.53-1b). This illusion was reported by Ernst Mach in the 19th century (Mach, 1865; Ratliff, 1965; Ross, Holt, & Johnstone, 1981) and now bears his name. Chevreul illusions (Cheveul, 1890; Ross et al., 1981; von Békésy, 1968) can be demonstrated with a luminance staircase that increases from step to step (Fig. III.53-2a–b). In the Craik-O'Brien-Cornsweet illusion (C-O-C; Cornsweet, 1970; Craik, 1940; O'Brien, 1958), concentric black and white rings imposed on a uniform surface change the (apparent) brightness of the entire circumscribed area (Fig. III.53-3a–b). In all of these illusions, the spatial distribution of perceived brightness diverges strikingly from the physical distribution of luminance. Lu and Sperling (1996) replaced the spatial variations of luminance in the stimuli that generate the three classic brightness illusions with spatial variations in texture contrast (Figures III.53-1c, III.53-2c, III.53-3c). The results are called second-order illusions because the theory for second-order processing involves two successive stages (Chubb & Sperling, 1989): first, a stage of rectification (a grossly nonlinear transformation) of luminance represented as positive and negative deviations of the luminance at a point from the overall mean luminance, and second, an analysis of the rectified deviations similar to the analysis of the original luminance stimuli.

The root of our second-order illusions is a random dynamic texture, the carrier, in which the luminance of each pixel is chosen randomly and independently and in which the expected luminance value of every pixel is the same. Imposed on this carrier texture there is a spatial modulation of contrast. The modulator of contrast in second-order textures serves the same role as a modulator of luminance in classical (first-order) patterns. However, in the second-order stimuli, only contrast modulation varies across space; the expected luminance is the same everywhere. To recover the modulator function from a first-order stimulus, we simply measure the luminance at each point and subtract the mean luminance. To recover the modulator function from a second-order texture stimulus, it is necessary to rectify the contrast values of each pixel of the stimulus. The (Weber) contrast of a pixel in an image is the luminance of that pixel minus the mean luminance of the image; that difference is then divided by the mean luminance of the image. A positive contrast indicates that a pixel is brighter than the mean brightness of the image, a negative contrast means that it is darker than the mean. Full-wave rectification is a monotonically increasing function of the absolute value of contrast (e.g., the absolute value itself, as in all the examples here) or the squared contrast. Except for random fluctuations, a random texture becomes equivalent to an ordinary luminance pattern upon absolute value rectification (Chubb & Sperling, 1989) and becomes accessible to standard linear analyses (matched linear filter followed by energy detection; e.g., Sperling, 1964).

Second-order Mach bands, Chevreul, and C-O-C illusions are shown in Figures III.53-1c, III.53-2c, and III.53-3c. In

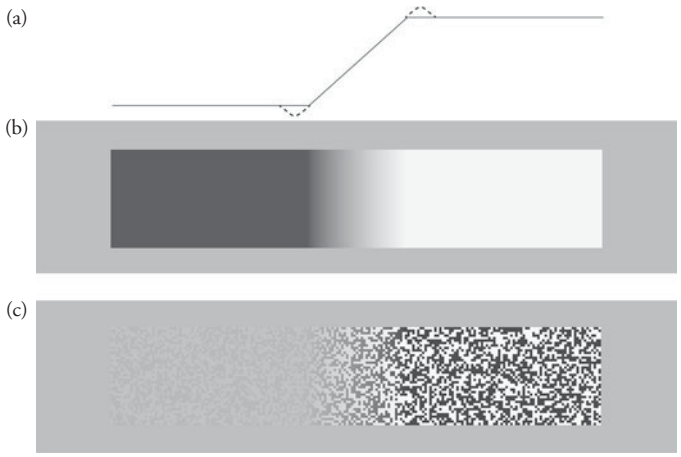


Figure III.53-1. (a) Luminance modulation function (luminance as a function of space) for a classical Mach band stimulus. The dotted curves illustrate the percept. (b) A first-order Mach band pattern. (c) A second-order Mach band pattern. Decreased brightness/contrast is perceived near the foot of the ramp, and increased brightness/contrast is perceived near the top of the ramp in (b) and (c). (Reproduced with permission of Z.-L. Lu and G. Sperling.)

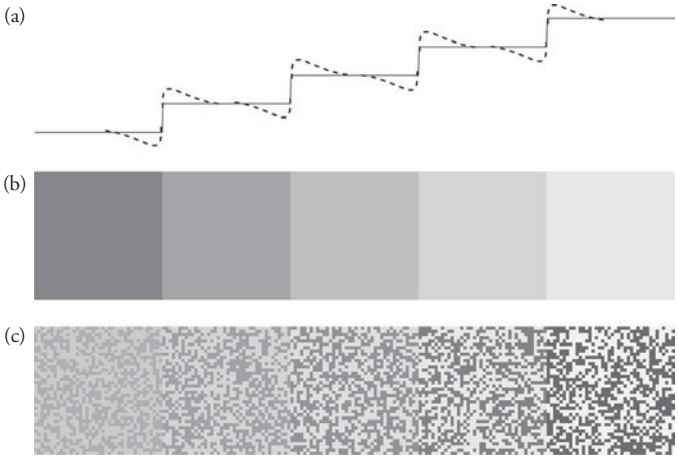


Figure III.53-2. (a) Luminance modulation function (luminance as a function of space) for a classical Chevreul stimulus. The dotted curves illustrate the percept. (b) A first-order Chevreul stimulus pattern. (c) A second-order Chevreul stimulus pattern. An illusory brightness/contrast valley is perceived at the foot of each step and an illusory peak brightness/contrast at the lip of the step in (b) and (c). (Reproduced with permission of Z.-L. Lu and G. Sperling.)

formal experiments, these stimuli are calibrated photometrically and psychophysically so that the modulator is invisible to first-order processes. Luminance calibration is not possible in a printed version. Also, the formal experiments involve dynamic stimuli, only a single frame of which is shown in the printed version. The magnitude of the second illusions depends on viewing distance—a close viewing distance works better.

When viewing the second-order Mach band patterns (Fig. III.53-1c), our observers reported that they perceived a low-contrast band near the low-contrast end of the ramp and a high-contrast band near the high-contrast end of the ramp. These subjective impressions were quite strong. A contrast-matching experiment using reference bars placed at six spatial locations, corresponding to the top and bottom of the ramp and the two plateaus, confirmed and quantified the general subjective impressions. We found Mach bands with generally similar magnitudes in luminance and in these second-order texture-contrast stimuli.

In the Chevreul illusion, a contrast modulator is made of monotonically increasing (or decreasing) steps. We generated both first-order (luminance) and second-order Chevreul stimuli (Fig. III.53-2). Each stimulus has five steps and all stimuli have the same overall spatial dimensions. When viewing the second-order texture-contrast and luminance Chevreul patterns, all six observers reported that they perceived low-contrast bands on the foot side of edges and high-contrast bands on the lip side of edges. These subjective impressions were quite strong. A nulling procedure confirmed and quantified the general subjective impressions: the magnitudes of the illusory Chevreul bands were estimated by the magnitudes of the incremental hills and valleys needed to cancel the illusion. The measured amplitudes were quite similar for the first-order luminance stimulus and for the second-order noise texture.

The stimuli that exhibit the C-O-C illusion are composed of textures whose modulator is a function only of radius in a polar coordinate system (Fig. III.53-3). In the experiment, modulator consists of four segments: (a) a constant

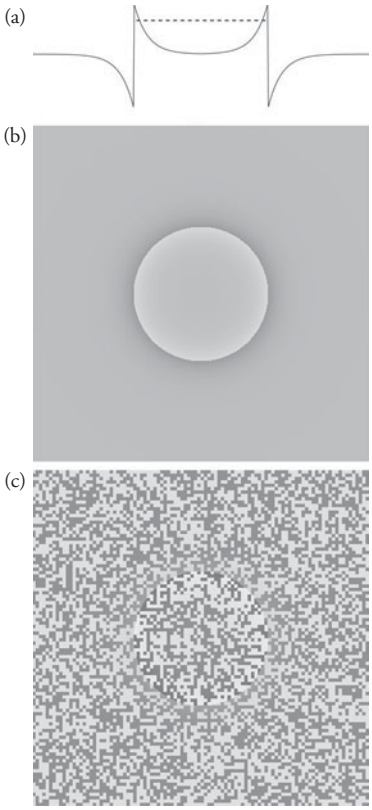


Figure III.53-3. (a) A luminance modulation function (luminance as a function of space) for a classical C-O-C stimulus. The dotted line illustrates the illusory percept. (b) A first-order C-O-C stimulus pattern. (c) A second-order C-O-C stimulus pattern. The illusion is an increase in apparent brightness/contrast of the entire central area in (b) and (c). (Reproduced with permission of Z.-L. Lu and G. Sperling.)

inner disk; (b) an inner ring with smaller than average values of the modulator; (c) an adjacent, concentric outer ring with larger values; and (d) the background with the same value as the inner disk. Seven comparison patterns, differing only in inner disk contrast values, were displayed simultaneously with the C-O-C patterns. Subjects were instructed first to look at the C-O-C pattern for 5 s and then to find and select the comparison stimulus whose inner disk contrast best matched the center contrast of the C-O-C pattern. All subjects—indeed everyone who has seen the full-wave C-O-C pattern—has invariably matched it to a comparison pattern with lower center contrast. The second-order C-O-C stimulus reliably produces an illusory reduction of contrast of magnitude comparable to the illusory contrast reduction in the first-order version.

Classical Mach bands, Chevreul, and C-O-C illusions are demonstrated in dynamic full-wave stimuli (i.e., stimuli made up of random carrier textures and modulators that, when subjected to absolute value rectification, become equivalent to the stimuli used in the original first-order illusions). None of these illusions is perceptible in half-wave stimuli (Lu & Sperling, 1996), that is, stimuli that are neutral to first-order (Fourier) and to second-order (full-wave) analyses but become equivalent to luminance stimuli after positive or negative half-wave rectification. Together these results indicate that the perceptual processes governing second-order spatial interactions, like those governing

second-order motion perception (Chubb & Sperling, 1989; Lu & Sperling, 1995; Solomon, Sperling, & Chubb, 1993), reflect full-wave (versus half-wave) rectification. Full-wave interaction was experimentally demonstrated as the modus of the second-order version of the simultaneous brightness contrast illusion (Chubb, Sperling, & Solomon, 1989). These results are experimental evidence in favor of the type of energy computations (energy is square-law full-wave rectification) proposed by Burr and Morrone (1994) to account for such illusions. In the spatial domain, as in motion, second-order processing of contrast-modulated stimuli, after full-wave rectification, is remarkably similar to first-order luminance processing.

REFERENCES

Burr, D. C., & Morrone, M. C. (1994). The role of features in structuring visual images. In CIBA Foundation Symposium (Ed.), *Higher order processing in the visual system* (pp. 129–146). New York: Wiley.

Cheveul, M. E. (1890). *The principles of harmony and contrast of colors* (C. Martel, Trans.). London: George Bell and Sons.

Chubb, C., & Sperling, G. (1989). Two motion perception mechanisms revealed through distance-driven reversal of apparent motion. *Proceedings of the National Academy of Sciences USA*, *86*(8), 2985–2989.

Chubb, C., Sperling, G., & Solomon, J. A. (1989). Texture interactions determine perceived contrast. *Proceedings of the National Academy of Sciences USA*, *86*(23), 9631–9635.

Cornsweet, T. N. (1970). *Visual perception*. Oxford: Academic Press.

Craik, K. J. W. (1940). *Visual adaptation* (PhD dissertation), Cambridge University, Cambridge.

Lu, Z.-L., & Sperling, G. (1995). The functional architecture of human visual motion perception. *Vision Research*, *35*(19), 2697–2722.

Lu, Z.-L., & Sperling, G. (1996). Second-order illusions: Mach bands, Chevreul, and Craik-O'Brien-Cornsweet. *Vision Research*, *36*(4), 559–572.

Mach, E. (1865). Über die Wirkung der räumlichen Vertheilung des Lichtreizes auf die Netzhaut. *Sitzungsberichte der mathematisch-naturwissenschaftlichen Classe der kaiserlichen Akademie der Wissenschaften*, *52*, 303–322.

O'Brien, V. (1958). Contour perception, illusion and reality. *Journal of the Optical Society of America A*, *48*(2), 112–119.

Ratliff, F. (1965). *Mach bands: Quantitative studies on neural networks in the retina*. Oxford: Holden-Day.

Ross, J., Holt, J., & Johnstone, J. (1981). High frequency limitations on Mach bands. *Vision Research*, *21*(7), 1165–1167.

Solomon, J. A., Sperling, G., & Chubb, C. (1993). The lateral inhibition of perceived contrast is indifferent to on-center/off-center segregation, but specific to orientation. *Vision Research*, *33*(18), 2671–2683.

Sperling, G. (1964). Linear theory and the psychophysics of flicker. *Documenta Ophthalmologica*, *18*(1), 3–15.

von Békésy, G. (1968). Mach-and Hering-type lateral inhibition in vision. *Vision Research*, *8*(12), 1483–1499.

Chapter 54
Vasarely's Nested Squares and the Alternating Brightness Star Illusion

Susana Martinez-Conde and Stephen L. Macknik

INTRODUCTION

Visual artists use their intuitive knowledge of perception—including illusory perception—to achieve specific results in their art. The optic art (op-art) movement, founded by Victor Vasarely, started in the 1960s in both Europe and the United States with a special focus on visual illusions as an art form per se, making explicit the long-standing implicit link between art, illusions, and visual science.

The term "op-art" was coined by *Time* magazine in 1964 and became very popular after the exhibition "The Responsive Eye," held at the Museum of Modern Art in New York, in 1965. Op-artists explored many aspects of perception, including illusory motion, geometrical relationships, impossible figures, and brightness, color, and shape illusions. Many of Vasarely's artworks exemplify how artists have sometimes preceded scientists in the discovery of fundamental principles of visual perception through the application of methodic—albeit intuitive—research techniques. One of Vasarely's best-known series, the nested squares paintings (Fig. III.54-1a–b), has given us significant insight into how corners, angles, curves, and line endings play a critical role in the appearance of shape and brightness in our brains.

VASARELY'S NESTED SQUARES ILLUSION

Vasarely's (1970) nested squares illusion shows that the corners of concentric squares, arranged in a gradient of increasing or decreasing luminance, generate illusory "folds," which appear more salient (i.e., either brighter or darker) than the adjacent flat (non-corner) regions of each individual square (Fig. III.54-1a–b). This classic illusion—though described often in the vision literature (Adelson, 1999; Hurvich, 1981; Kaiser, 1996; McArthur & Moulden, 1999; Morgan, 1996)—was not quantified or tested systematically, with computational, psychophysical, or neurophysiological techniques, until recently. We and our colleagues set out to determine its underlying mechanisms and to use our findings to develop new illusions and principles of visual perception (Troncoso, Macknik, & Martinez-Conde, 2005, 2009; Troncoso et al., 2007).

THE ALTERNATING BRIGHTNESS STAR ILLUSION

In the classic Vasarely's nested squares artwork, corner angles are always 90º (Figure III.54-1a–b). To determine the significance of angle sharpness and sign (i.e., concavity or convexity) on the brightness of the illusory folds, we varied the angles of the corners. By doing this, we discovered a novel illusion that we called the "alternating brightness star" (ABS; Martinez-Conde & Macknik, 2001; see Figures III.54-1c–d and III.54-2). The ABS—composed of concentric stars arranged in gradients of decreasing or increasing luminance—produces alternating bright *and* dark illusory folds, depending on whether the angles of the corners are concave or convex (Fig. III.54-1c–d). We called this effect "corner angle brightness reversal" (Troncoso et al., 2005). The ABS also shows that the magnitude of the perceived contrast varies with corner angle: acute angles are perceived as high contrast (Fig. III.54-1c–d, top), and shallow angles exhibit lower contrast (Fig. III.54-1c–d, bottom). We called this effect "corner angle salience variation." See Interactive Figure III.54-1 on the Companion Website.

Physical luminance (Fig. III.54-2e) and optical blurring due to the physical limitations of the eye (Figures III.54-2c and III.54-2f) do not explain the ABS. Instead, the illusion is due to changes in local contrast (Figures III.54-2d and III.54-2g). Indeed, when the ABS is filtered with a difference of Gaussians (DOG) filter, the output qualitatively matches the perception of the illusion (Figures III.54-2d and III.54-2g). The DOG filter simulates the processes conducted by the center-surround receptive fields of neurons in the retina and the lateral geniculate nucleus. Center-surround receptive fields consist of two concentric regions that are subtracted from each other to provide a measure of contrast between the two corresponding areas in the visual field.

PERCEPTUAL QUANTIFICATION OF VASARELY'S NESTED SQUARES ILLUSION AND THE ABS

In collaboration with our colleague Xoana G. Troncoso, then at the Barrow Neurological Institute, we set out to quantify the perceptual strength of the nested squares and

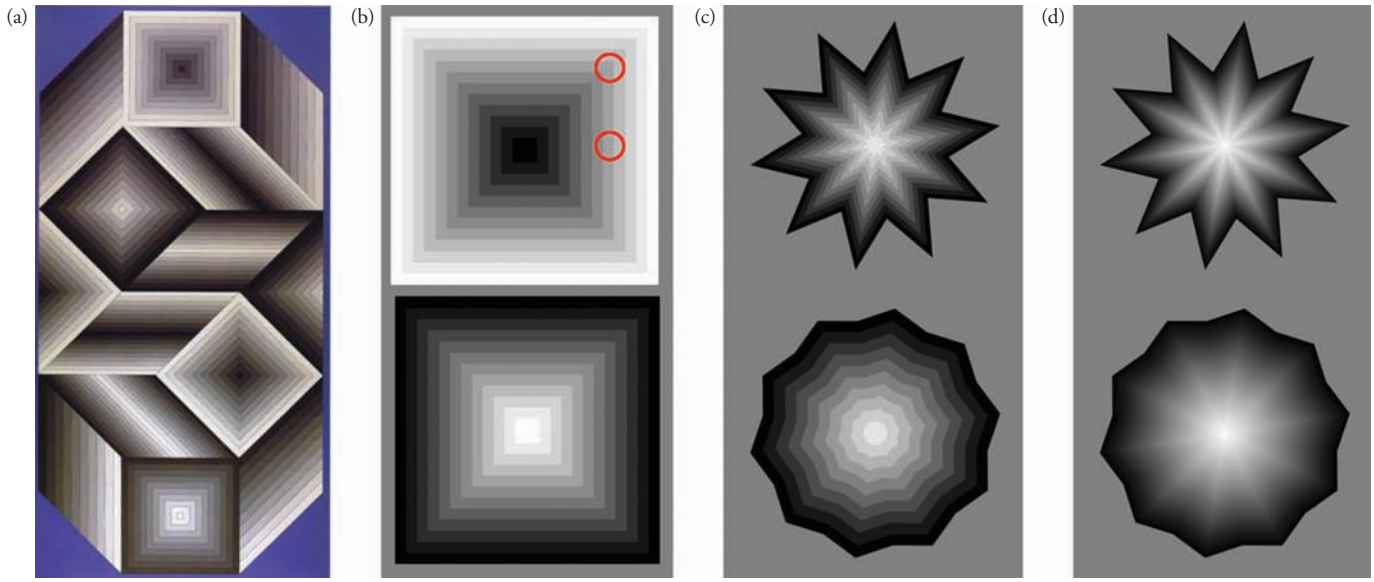


Figure III.54-1. Vasarely's nested squares and ABS illusions. (a) Vasarely's "Utem" (1981). Note the four sets of nested squares. The two nested squares of decreasing luminance (from the center to the outside) have bright illusory diagonals, whereas the two nested squares of increasing luminance (from the center to the outside) have dark illusory diagonals. The physical luminance of each individual square remains constant at all points, but the corners appear perceptually more salient than the squares' straight edges, forming illusory X shapes that seem to irradiate from the very center of each set of squares. (Reproduced with kind permission by the Vasarely Foundation from Vasarely, 1982.) (b) Nested squares illusion based in Vasarely's "Arcturus" (Vasarely, 1970). Top: The stimulus is made out of multiple concentric squares of increasing luminance (from the center to the outside). The two red circles indicate two regions that appear different in brightness. The area inside the upper red circle has higher average luminance—but appears perceptually darker—than the region inside the lower circle. Bottom: Nested square stimulus, with a gradient of decreasing luminance (from the center to the outside). (From Troncoso et al., 2005.) (c) and (d) The ABS illusion (Martinez-Conde & Macknik, 2001). The stimulus is made of concentric stars of graded luminance. In the examples illustrated, the innermost star is white and the outermost star is black. The illusory corner-folds that radiate from the center appear as light or dark depending on the polarity of the corner angle (i.e., the corner angle brightness reversal effect). Moreover, the illusory folds appear more salient with sharp corners (top stars) and less salient with shallow corners (bottom stars; i.e., the corner angle brightness variation effect). All illusory folds are physically equal to each other in luminance, however. (c) The gradient from the center to the outside has 10 luminance steps, making the individual stars forming the polygonal constructs easy to identify. (d) The gradient from the center to the outside has 100 luminance steps. (From Troncoso et al., 2009.)

ABS illusions (Fig. III.54-3). To do this, we used a psychophysical two-alternative forced-choice brightness discrimination task. Human subjects compared the brightness of an illusory fold to the brightness of a standard stimulus. The illusory fold was made out of corners of seven possible angles, and the point of comparison always had a luminance of 50% grey. The standard stimulus had 11 difference possible luminance levels (Fig. III.54-3a–d). The perceived brightness of the corner varied parametrically with corner angle, with sharp angles leading to strong illusory percepts and shallow angles leading to weak percepts (corner angle salience variation; Fig. III.54-3e–f; Troncoso et al., 2005).

CENTER-SURROUND SIMULATIONS

A recurrent theme in Vasarely's art is the generation of illusory percepts from deflections or discontinuities in edges, such as corners, curves, and terminators. Note that the zebra in Figure III.54-4a is solely constructed from small deflections in the parallel lines that run diagonally across the surface of the image. The zebra has no actual contour: our brain samples the most salient parts of the image (i.e., the points of maximum curvature) and connects them together to create the percept of a galloping zebra. When the image is convolved with a center-surround DOG filter (Enroth-Cugell & Robson, 1966; Rodieck, 1965; i.e., to enhance local contrast; see later discussion), the contour of the zebra pops out (Fig. III.54-4b).

In 1981 Hurvich proposed that center-surround receptive fields could account for the nested squares illusion; that is, that the contrast between the center and the surround regions of the receptive field would be stronger along the corner gradients than along the edge gradients in a nested square pattern, resulting in increased perceptual salience at the corners. Grossberg and Todorović (1988) later simulated the responses from geniculate and primary visual cortex neurons to various visual stimuli and found enhancements for 90° corners. They proposed that such increased responses could contribute to the perception of Vasarely's nested squares illusion as described by Hurvich (Grossberg & Todorović, 1988).

We wondered whether an extension of Hurvich's (1981) model, now applied to corner gradients of all angles, could explain the psychophysical results described previously. To test this possibility, we modeled center-surround receptive fields as DOG filters using realistic physiological parameters (see Troncoso et al., 2005, for details). Figure III.54-4c shows that convolving a DOG filter with the stimuli from the psychophysical experiments reproduces

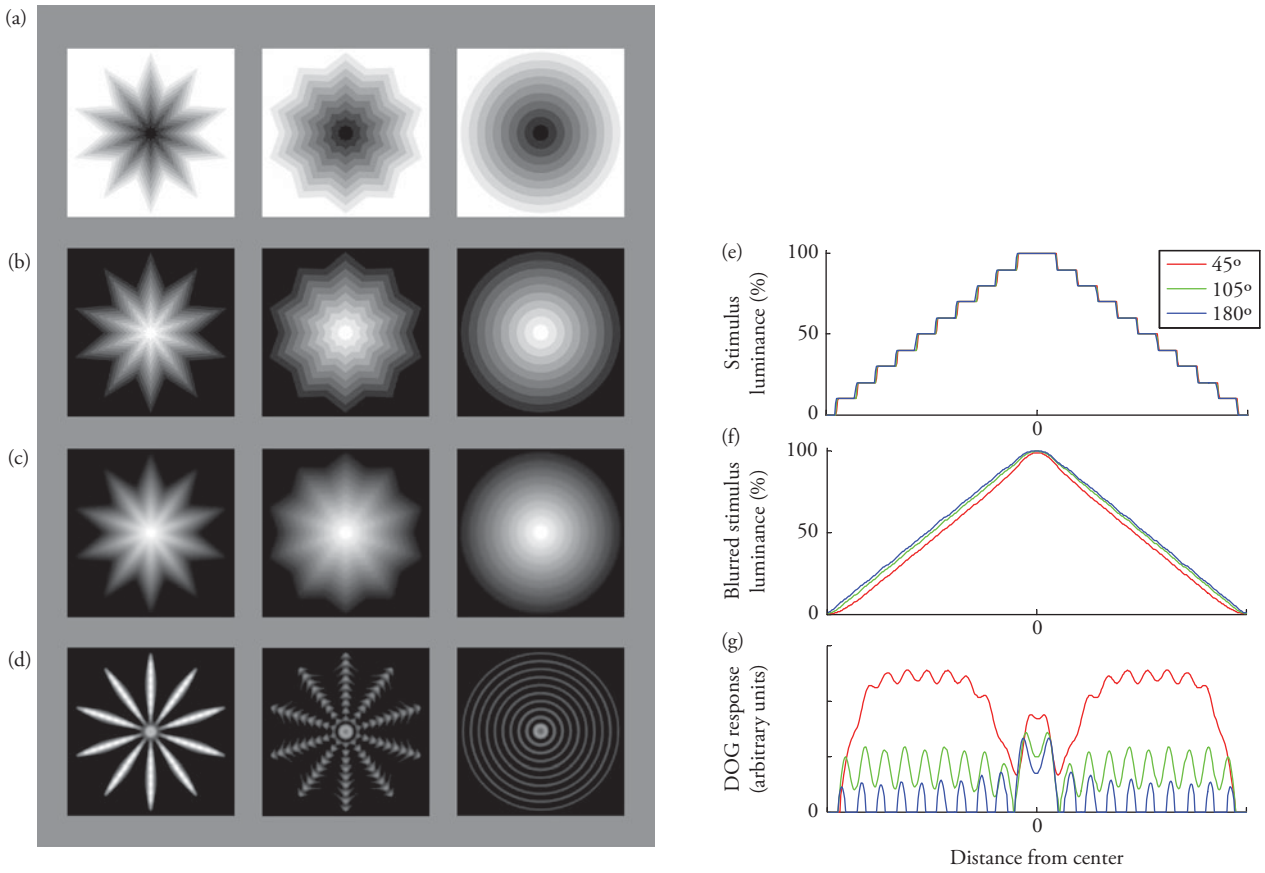


Figure III.54-2. The ABS illusion (Martinez-Conde & Macknik, 2001; Troncoso et al., 2005). (a) and (b) The gradient from the center to the outside of each set of stars has 10 luminance steps. In (a), the innermost star is black and the outermost star is white. In (b), the innermost star is white and the outermost star is black. The illusory folds that radiate from the center of each set of stars appear light or dark depending on the polarity of the corner angle and the direction of the luminance gradient (i.e., corner angle brightness reversal effect). However, all illusory folds are physically equal to each other in luminance. Sharp corner angles (45° column) result in stronger illusory diagonals. Shallow corner angles produce weaker (105° column) or absent (180° column) illusory diagonals. (c) Blurred version of the stars in (b) (Gaussian filtering). (d) Result of filtering the stimuli in (b) with a DOG filter, simulating the output of center-surround receptive fields. (e) Luminance profile along one of the arms of the star for the stimuli in (b). The illusory folds have the same physical luminance in the three-angle conditions (the three lines overlap). (f) Luminance profile along one of the arms of the star for the stimuli in (c). The illusory folds cannot be explained by blurring: blurring causes sharp corners to have lower luminance than shallow corners, contrary to our perception of the illusion. (g) Luminance profile along one of the arms of the star for the stimuli in (d). The center-surround simulation qualitatively matches our perception of the illusion: the predicted response is highest for the sharpest corner angle and decreases as the corner angle becomes shallower. (Modified from Troncoso et al., 2007.)

qualitatively the corner angle salience variation effect (i.e., the strength of the illusion varies parametrically with the angle of the corner, with sharp angles producing stronger outputs).

Following from this, we proposed a further extension of Hurvich's (1981) hypothesis into a general model of corner processing in the early visual system, as the geometry of retinogeniculate center-surround receptive fields makes them, in general, better corner detectors than edge detectors (Fig. III.54-4d). Because neurons in higher visual areas (such as the primary visual cortex) integrate visual information from retinogeniculate levels (Alonso & Martinez, 1998; Ferster, 1986; Ferster, Chung, & Wheat, 1996; Hubel & Wiesel, 1962; Martinez & Alonso, 2001; Reid & Alonso, 1995; Tanaka, 1983), cortical receptive fields should also be better tuned to corner detection than to edge detection. Thus corner and curve detection may first occur within the lowest levels of the visual system, even if such curves and corners play an important role in object recognition and shape processing in higher visual areas.

BRAIN ACTIVATION VARIES PARAMETRICALLY WITH CORNER ANGLE

Our psychophysical studies led us to the hypothesis that the nested squares and ABS illusions arise because of the interaction between surface corners and the shape of receptive fields in early visual areas. Sharp corners may create hotspots of high local contrast due to processing by center-surround and other early receptive fields (Fig. III.54-4d).

If this hypothesis is correct, early visual neurons should respond powerfully to sharp corners and curved portions of surface edges. Indeed, a fundamental role of early visual neurons may be to localize the discontinuities along the

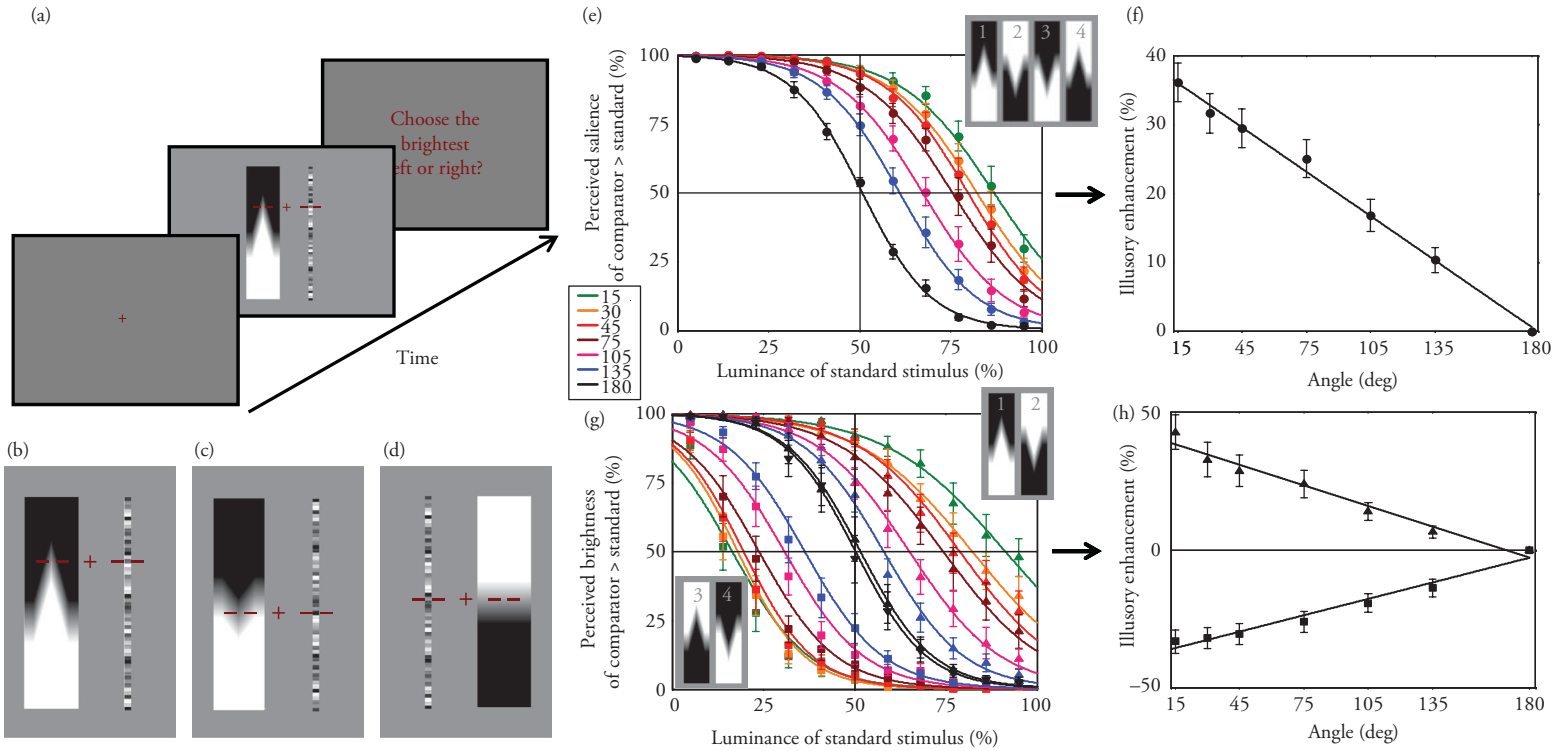


Figure III.54-3. Perceptual quantification of the nested squares and ABS illusions. (a) Monitor display during time course of a single experimental trial. (b), (c), and (d) Three different stimuli presentations of the brightness discrimination task (out of 572 possible conditions; see Troncoso et al., 2005). (e) Psychometric functions for the different corner angles are plotted in different colors. Gradient and angle polarities are collapsed: that is, we averaged together the conditions where the illusory folds of the comparator looked bright and the conditions where they looked dark (Insert: Illustrations of the four collapsed conditions for a 30° angle). (f) Illusory enhancement of the points of subjective equality (PSEs) with respect to the control condition (180° gradient) for the different comparator corner angles. The illusory enhancement decreases linearly as the angle of the comparator corner gradient becomes shallower (corner angle salience variation effect). (g) Same data as in (e) but the conditions in which the illusory fold looked bright (triangles) and the conditions in which the illusory fold looked dark (squares) are presented separately (see Inserts). The PSEs for conditions 1 and 2 (see insert) fall above the point of 50% luminance in the standard stimulus. The PSEs for conditions 3 and 4 (insert) fall below the point of 50% luminance in the standard stimulus (corner angle brightness reversal effect). (h) Illusory enhancement of brightness (triangles) and darkness (squares) perception, as a function of corner angle. Error bars in (e), (f), (g), and (h) represent the ± standard error of the mean for all subjects in each condition. (From Troncoso et al., 2005.)

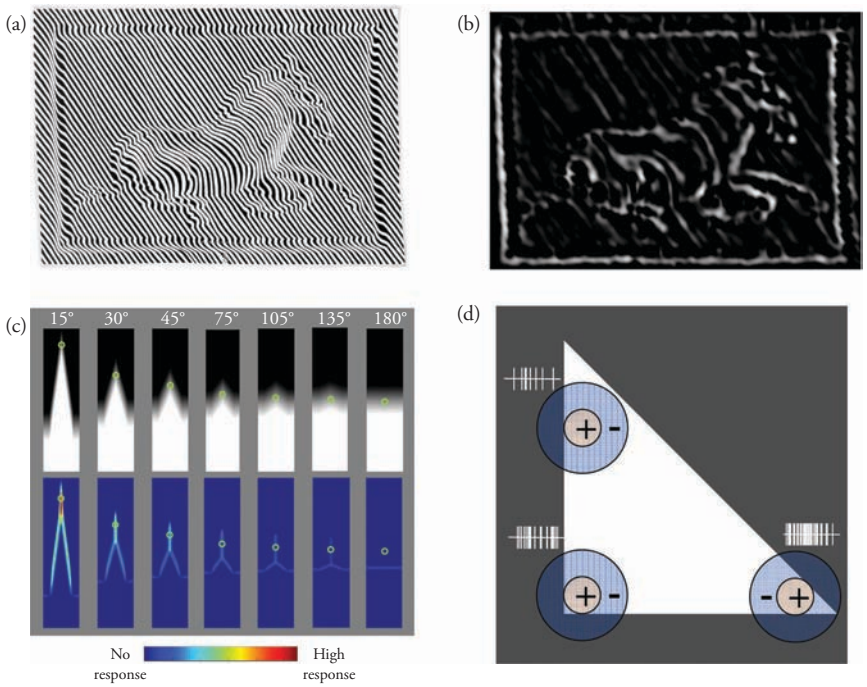


Figure III.54-4. Center-surround receptive field responses to corners. (a) Vasarely's "Zebra" (1938), from Aknai and Sarkany (2000). The zebra has no actual contour: our brain samples the points of maximum curvature (i.e., the deflections in the diagonal lines) and connects them together to create the percept of a galloping zebra. (Reproduced with kind permission from the Vasarely Foundation.) (b) The image in (a) has been convolved with a center-surround filter. This simple transformation, which is also performed by the center-surround receptive fields of the visual system, makes the contour of the zebra highly salient. (From Troncoso et al., 2009.) (c) Computational simulations of the ABS illusion with a DOG filter. The filter parameters were chosen to match physiological center-surround receptive fields at the eccentricity used in the psychophysical experiments (3º). Top: Examples of corner-gradient stimuli analyzed in the simulations. The green circles mark the point of 50% luminance. Bottom: Convolving the DOG filter with the corner-gradient stimuli simulates the output of an array of center-surround neurons. The green circles indicate the responses of the model at the point of 50% luminance on the actual gradient. (From Troncoso et al., 2005.) (d) Generalized model of corner processing. On-center receptive fields are placed over one edge and two corners of a white triangle. The center of the receptive field over the edge is well stimulated by light, but most of the surround also falls in the light region, so the response of the neuron is partially inhibited. The center of the receptive field over the 90º corner is also stimulated by light and most of the surround falls in the dark area. This is a more optimal stimulus than in the previous location and leads to a stronger neural response. The receptive field over the 45º corner receives even more optimal contrast between center and surround, leading to an even stronger response. The spiking responses depicted in the cartoon are hypothetical. (From Troncoso et al., 2005.)

edges of surfaces. If so, all early visual areas should show greater responses to sharp corners than to shallow corners. Conversely, if corner processing is constrained to certain areas of the visual cortex (as predicted by previous studies; see Hubel & Wiesel, 1965; Ito & Komatsu, 2004; Pasupathy & Connor, 1999), only a subset of visual areas would show greater responses to sharp than to shallow corners. To address this, we teamed up with Xoana G. Troncoso, Alexander Schlegel, and Jorge Otero-Millan, then at the Barrow Neurological Institute, and Peter Tse, Gideon Caplovitz, and Po-Jang Hsieh, then at Dartmouth College, to explore the neural correlates of the ABS, particularly the corner angle salience variation effect in the human visual cortex, via functional magnetic resonance imaging (fMRI; Troncoso et al., 2007).

Before this research, no neuroimaging or electrophysiological studies had determined the contribution of angle sharpness to corner processing in any species. We presented ABS with equivalent angles to those used in the earlier psychophysical experiments (Troncoso et al., 2005), while acquiring fMRI data at the Dartmouth Brain Imaging Center. We localized the retinotopic cortical regions that modulated their activity in correlation with the angle of the corner. The BOLD signal correlates of the ABS matched the psychophysical results: BOLD signal responses were stronger for sharp angles than for shallow angles, suggesting that the corresponding neurons were more active in response to sharp corners than to shallow corners. Further, all individual retinotopic areas displayed a parametric relationship between corner angle and strength of BOLD signal (Fig. III.54-5). This suggests that corner and angle processing are generalized properties of the early visual system and not solely the function of a small subset of circuits in cortical areas beyond the primary visual cortex (as previous models of corner processing in the brain had predicted). The results were consistent with our proposal (Troncoso et al., 2005) that most visual neurons respond more strongly to sharp than to shallow corners, including neurons with center-surround receptive fields in the retina and lateral geniculate nucleus, and ruled out the possibility that all corner processing takes place in a specific retinotopic cortical region (or subset of retinotopic regions). These results offered the first physiological correlates of the ABS and the nested squares illusion, and suggested that sharp corners are more powerful stimuli

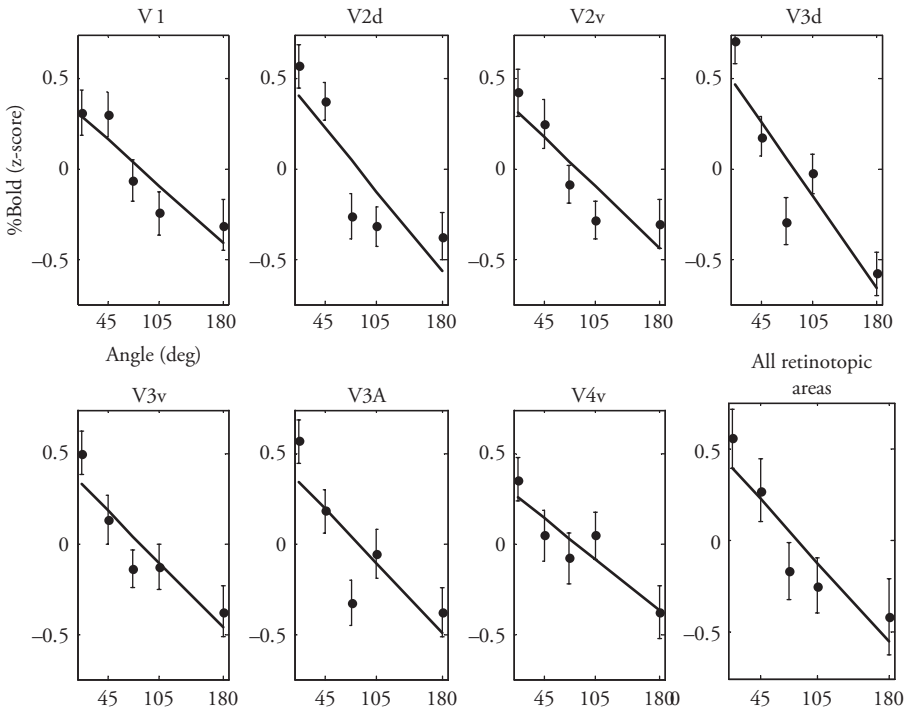


Figure III.54-5. BOLD signal as a function of corner angle in human retinotopic areas. Each data point was obtained by first averaging the responses between 5 and 20 s after stimulus onset, for each subject's hemisphere, and then computing the *z* score across the corner angle conditions ($n = 24$). Regression lines were calculated with the least-squares fit. Error bars represent structural equation modeling between subject hemispheres. (From Troncoso et al., 2007.)

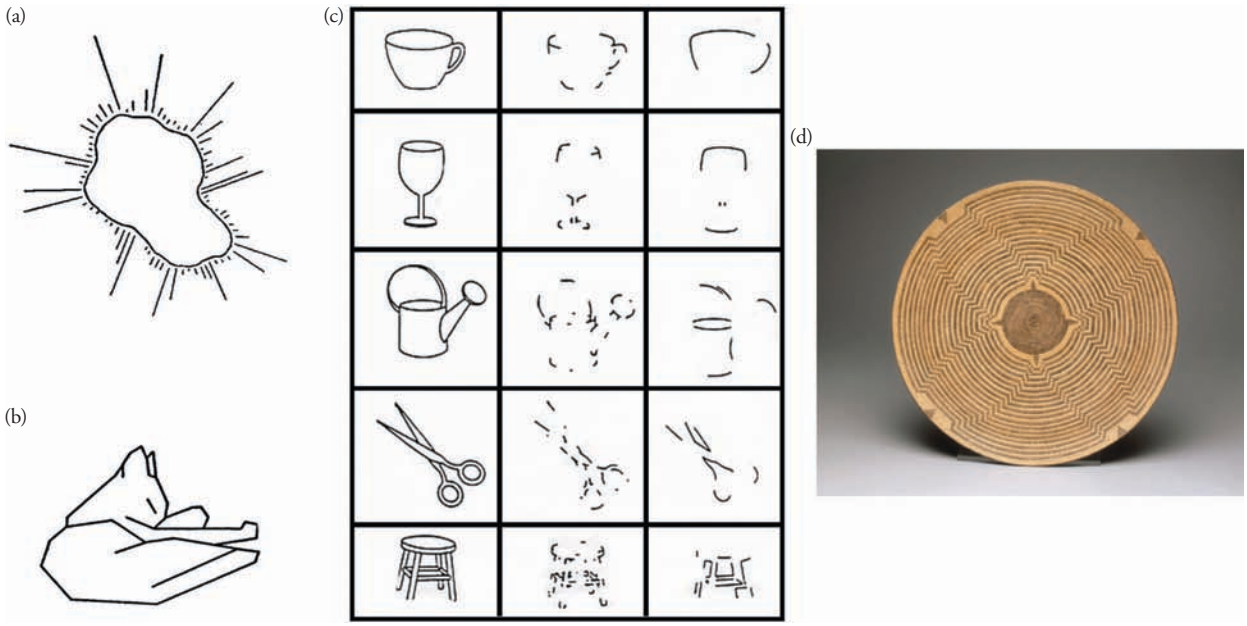


Figure III.54-6. Corners and curves in shape and depth perception. (a) and (b) Attneave's corner and curvature studies. (a) Subjects were asked to choose the 10 points of the blob that would be most useful for reconstructing its shape. The radiating lines indicate the frequency with which each point was chosen. Sharp curves and corners were chosen much more often than shallow curves and straight edges. (b) This drawing was made by choosing the 38 points of maximum curvature from the contour of a sleeping cat and connecting them by straight lines. The shape of the animal is very easily recognizable. (From Attneave, 1954.) (c) Biederman's studies on object recognition. The left column shows line drawings of five common objects. When the high-curvature sections of the contour are illustrated (middle column), the objects are still easily recognizable. But when only the straight edges and shallow curves are depicted (right column), the objects are very difficult to recognize (even though as much contour is represented as in the middle column). (From Biederman, 1987.) (d) Pima (Akimel O'odham) basket. The flat tray induces the perception of a windmill whose wedges recede in depth relative to the ground (Shapley & Maertens, 2008). (Image from the Charles and Valerie Diker Collection [Kuspit et al., 2004]; reproduced with the permission of Charles Diker.)

than shallow corners throughout the human occipital cortex.

CORNER PERCEPTION AND THE REDUNDANCY-REDUCING HYPOTHESIS

The information transmitted by our visual system is constrained by physical limitations, such as the relatively small number of axons available in the optic nerve. To some extent, our visual system overcomes these limitations by extracting, emphasizing, and processing nonredundant visual features. In 1961 Barlow proposed that the brain recodes visual data "so that their redundancy is reduced but comparatively little information is lost," an idea known as the redundancy-reducing hypothesis (Barlow, 1961, 1989). The redundancy-reducing hypothesis has been invoked as an explanation for why neurons at the early levels of the visual system are suited to perform "edge detection," or "contour extraction." However, redundancy reduction is not necessarily constrained to edges but rather should theoretically apply to any feature in the visual scene (Rao, Olshausen, & Lewicki, 2002). Fred Attneave (1954) proposed that "points of maximum curvature" (i.e. discontinuities in edges, such as curves, angles, and corners—any point at which straight lines are deflected) are less redundant than edges themselves and thus contain more information. If points of high curvature are less redundant than points of low curvature, then sharp corners should also be less redundant than shallow corners. This led to our prediction that sharp corners should appear more salient, and generate stronger physiological responses, than shallow corners. Figure III.54-6a–b illustrates examples from Attneave's and Biederman's (1987) studies on the importance of corners for shape perception and object recognition.

BEYOND BRIGHTNESS PERCEPTION: COLOR AND DEPTH EFFECTS IN THE NESTED SQUARES AND ABS ILLUSIONS

In recent years, a handful of studies have started to investigate nonbrightness effects in the nested squares and ABS illusions. Yucht, Beyil, Einav, and Spitzer (2010) presented chromatic versions of the nested squares and ABS illusions where the concentric angles formed a chromatic—though equiluminant—gradient. Psychophysical testing showed that the perceived colors of the illusory folds were complementary to the colors of the concentric corners. Moreover, sharper corner angles generated stronger illusory folds, in agreement with the corner angle salience variation effect (Troncoso et al., 2005). Shapley and Maertens (2008) studied the three-dimensional component of the illusory folds perceived in the nested squares and similar illusions by Victor Vasarely and other artists, including the American Indian basketry artists of the Akimel O'odham (Pima) tribe (Fig. III.54-6c). They found that the amount of perceived depth depended on the steepness of the corner angle, also consistent with the corner angle salience variation effect (Troncoso et al., 2005).

CONCLUSIONS

Our combined results from human psychophysics experiments, human brain imaging, and computational modeling suggest that deflections or discontinuities in edges, such as corners, curvature, and terminating line endings, may be first processed by center-surround receptive fields. The antagonism between receptive field center and surround results in sharp corners, curves, or endpoints being more potent stimuli than shallow corners/curves or flat edges (Troncoso et al., 2005). Thus changes in BOLD signals with corner angles may be primarily driven by changes in local contrast at the corners (Troncoso et al., 2007).

This hypothesis in no way rules out a critical role for later cortical areas in more complex processing of corner angles. For instance, we expect that specific orientations of corner angles must be processed cortically, given that the first orientation-selective cells are cortical.

In summary, our results offer insights into the neural mechanisms responsible for corner processing and also provide new data regarding the stage of the visual hierarchy in which corner processing may arise. Future studies should determine whether responses from single neurons in the early visual system affect our perception of the nested squares and ABS illusions (Troncoso et al., 2005). Single-neuron studies may further test the hypothesis that the physiological and perceptual effects described here are due to the interaction of corners and receptive field organization.

ACKNOWLEDGMENTS

We thank Jorge Otero-Millan and Xoana G. Troncoso for help with the figures. This study was supported by the Barrow Neurological Foundation (SLM and SMC) and National Science Foundation Awards 0852636 and 1153786 (SMC).

REFERENCES

Adelson, E. H. (1999). Lightness perception and lightness illusions. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 339–351). Cambridge, MA: MIT Press.

Alonso, J. M., & Martinez, L. M. (1998). Functional connectivity between simple cells and complex cells in cat striate cortex. *Nature Neuroscience*, *1*(5), 395–403.

Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, *61*(3), 183–193.

Barlow, H. B. (1961). Possible principles underlying the transformation of sensory messages. In W. A. Rosenblith (Ed.), *Sensory communication* (pp. 217–234). Cambridge, MA: MIT Press.

Barlow, H. B. (1989). Unsupervised learning. *Neural Computation*, *1*, 295–311.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*(2), 115–147.

Enroth-Cugell, C., & Robson, J. G. (1966). The contrast sensitivity of retinal ganglion cells of the cat. *Journal of Physiology*, *187*, 517–552.

Ferster, D. (1986). Orientation selectivity of synaptic potentials in neurons of cat primary visual cortex. *Journal of Neuroscience*, *6*, 1284–1301.

Ferster, D., Chung, S., & Wheat, H. (1996). Orientation selectivity of thalamic input to simple cells of cat visual cortex. *Nature*, *380*, 249–252.

Grossberg S., &Todorović, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena. *Perception & Psychophysics*, *43*, 241–277.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, *160*, 106–154.

Hubel, D. H., & Wiesel, T. N. (1965). Receptive fields and functional architecture in two non-striate visual areas (18 and 19) of the cat. *Journal of Neurophysiology*, *28*, 229–289.

Hurvich, L. M. (1981). *Color vision.* Sunderland, MA: Sinauer Associates.

Ito, M., & Komatsu, H. (2004). Representation of angles embedded within contour stimuli in area V2 of macaque monkeys. *Journal of Neuroscience*, *24*(13), 3313–3324.

Kaiser, P. K. (Prod.). (1996). *The joy of visual perception: A web book*. http://www.yorku.ca/eye/

Kuspit, D., Bernstein, B., McMaster, G., Dubin, M., Ash-Milby, K. E. (2004). *First American Art: The Charles and Valerie Diker Collection of American Indian Art*. National Museum of the American Indian.

Maertens, M., & Shapley, R. (2008). Local determinants of contour interpolation. *Journal of Vision*, *8*(7), 3, 1–11.

Martinez, L. M., & Alonso, J. M. (2001). Construction of complex receptive fields in cat primary visual cortex. *Neuron*, *32*(3), 515–525.

Martinez-Conde, S., & Macknik, S. L. (2001). *Junctions are the most salient visual features in the early visual system*. November 10–15. Paper presented at the Society for Neuroscience, San Diego, CA.

McArthur, J. A., & Moulden, B. (1999). A two-dimensional model of brightness perception based on spatial filtering consistent with retinal processing. *Vision Research*, *39*(6), 1199–1219.

Morgan, M. J. (1996). Visual illusions. In V. Bruce (Ed.), *Unsolved mysteries of the mind: Tutorial essays in cognition* (pp. 29–58). New York: Psychology Press.

Pasupathy, A., & Connor, C. E. (1999). Responses to contour features in macaque area V4. *Journal of Neurophysiology*, *82*(5), 2490–2502.

Rao, R. P. N., Olshausen, B. A., & Lewicki, M. S. (2002). *Probabilistic models of the brain: Perception and neural function*. Cambridge, MA: MIT Press.

Reid, R. C., & Alonso, J. M. (1995). Specificity of monosynaptic connections from thalamus to visual cortex. *Nature*, *378*(6554), 281–284.

Rodieck, R. W. (1965). Quantitative analysis of cat retinal ganglion cell response to visual stimuli. *Vision Research*, *5*, 583–601.

Shapley, R., & Maertens, M. (2008). Angle alignment evokes perceived depth and illusory surfaces. *Perception*, *37*(10), 1471–1487.

Tanaka, K. (1983). Cross-correlation analysis of geniculostriate neuronal relationships in cats. *Journal of Neurophysiology*, *49*, 1303–1318.

Troncoso, X. G., Macknik, S. L., & Martinez-Conde, S. (2005). Novel visual illusions related to Vasarely's "nested squares" show that corner salience varies with corner angle *Perception*, *34*(4), 409–420.

Troncoso, X. G., Macknik, S. L., & Martinez-Conde, S. (2009). Corner salience varies linearly with corner angle during flicker-augmented contrast: A general principle of corner perception based on Vasarely's artworks. *Spatial Vision*, *22*(3), 211–224.

Troncoso, X. G., Tse, P. U., Macknik, S. L., Caplovitz, G. P., Hsieh, P. J., Schlegel, A., ... Martinez-Conde, S. (2007). BOLD activation varies parametrically with corner angle throughout human retinotopic cortex. *Perception*, *36*(6), 808–820.

Tsofe, A., Yucht, Y., Beyil, J., Einav, S., & Spitzer, H. (2010). Chromatic Vasarely effect. *Vision Research*, *50*(22), 2284–2294.

Vasarely, V. (1965). *Vasarely I.* Neuchâtel, Switzerland: Editions du Griffon.

Vasarely, V. (1970). *Vasarely II* (H. Chevalier, Trans.). Neuchâtel, Switzerland: Editions du Griffon.

Vasarely, V. (1982). *Gea*. Paris: Editions Hervas.

Chapter 55
Grating Induction

Mark E. McCourt and Barbara Blakeslee

INTRODUCTION

Grating induction (McCourt, 1982) is a brightness/lightness illusion in which a sinewave luminance grating induces the appearance of a counterphase spatial variation (i.e., a grating) in a homogeneous test field oriented orthogonally to the inducing grating; (Fig. III.55-1e). Although extensively studied by a few, the illusion has been largely ignored because of the difficulty of explaining the induced structure (brightness/lightness gradients) within the test field using currently popular (nonfiltering-based) models. Other older brightness/lightness effects commonly acknowledged to possess gradient structure have also been largely ignored of late (but see Geier & Hudak, 2011) and include Mach bands (Ratliff, 1965), the Chevreul illusion (Chevreul, 1890; Geier & Hudak, 2011), the Hermann grid illusion (Hermann, 1870; Spillmann, 1994), and linear gradient induction (Blakeslee & McCourt, 2005; Frisby, 1980; Kingdom, 1999; Moulden & Kingdom, 1991; Shapley & Reid, 1985). In addition to possessing a gradient structure, grating induction is also a very powerful brightness/lightness induction effect. As measured using a canceling technique, where a real luminance grating is added to the test field in opposite spatial phase to the induced grating, thus canceling it, grating induction magnitude can exceed 90% of inducing grating contrast (Foley & McCourt, 1985; McCourt, 1982; McCourt & Blakeslee, 1994). As measured using a matching procedure the contrast of induced gratings asymptotes at about 30% (McCourt & Blakeslee, 1994).

McCourt (1982) and Foley and McCourt (1985) found that the strength of grating induction depends on a number of factors, including (a) the height of the homogeneous test field, where induction magnitude increases with decreasing test field height; (b) the spatial frequency of the inducing grating, where induction magnitude is strongest at low spatial frequencies; (c) the height of the inducing grating, where induction magnitude increases (and eventually asymptotes) with increasing inducing grating height; and (d) the contrast of the inducing grating, where induction magnitude increases with increasing contrast. In addition, Foley and McCourt discovered that the magnitude of grating induction was determined by the product of inducing grating frequency and test field height, that is, by the number of inducing grating cycles that fit into the test field. The practical implication of this rule is that induction magnitude is roughly independent of viewing distance. Its larger, theoretical significance is that this behavior suggests to some that multiscale spatial filtering might play a causal role in the grating induction effect (Foley & McCourt, 1985; Kingdom & Moulden, 1992; Moulden & Kingdom, 1991).

The idea that grating induction might result from spatial filtering at relatively early levels in the visual system is further supported by the finding that induced gratings produced distortions of both the contrast and phase of physical luminance gratings added to them (McCourt & Blakeslee, 1994) and acted as pedestals to facilitate the detection of luminance gratings to a degree identical to that found for a real grating with the same spatial characteristics and perceived contrast (McCourt & Kingdom, 1996). The fact that the induced brightness/lightness variations can act as nearly perfect spatial metamers of real luminance variations is most parsimoniously accounted for by the idea that the same mechanisms that transduce real luminance variations (i.e., contrast) are transducing the illusory brightness/lightness variations as well. These mechanisms are generally understood to be retinal and cortical neurons whose receptive fields perform bandpass spatial filtering operations on the retinal luminance distribution (DeValois, Albrecht, & Thorell, 1982; DeValois & DeValois, 1988; Geisler, Albrecht, & Crane, 2007; Wilson & Wilkinson, 2003).

Seeking to explore the taxonomic relationship between grating induction and the much more frequently studied classical simultaneous contrast effect, Blakeslee and McCourt (1997) used point-by-point matching to obtain brightness/lightness profiles for a series of grating induction stimuli in which test field width transitioned smoothly from the two test patch configuration of simultaneous contrast to the single elongated test field of the grating induction configuration (Fig. III.55-1a–e).

These matching profiles (Fig. III.55-1f–j) revealed that horizontal induced brightness/lightness structure was present in the test fields of all of these stimuli (i.e., none of the test fields appeared homogeneous) and that both the structure and the magnitude of these brightness/lightness variations, as well as those in the novel grating induction configurations of Zaidi (1989), were parsimoniously

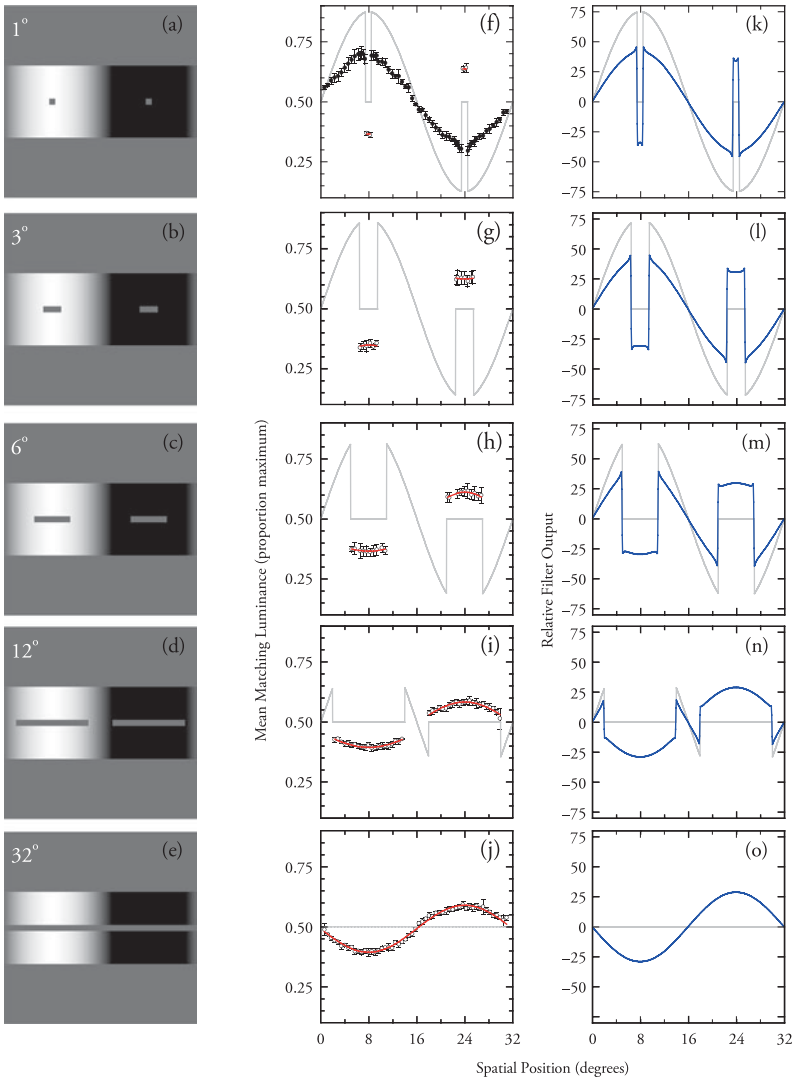


Figure III.55-1. Stimulus series (a) through (e) in which test field width transitions smoothly from the two test patch configuration of simultaneous brightness/lightness contrast (a) to the single elongated test field of the grating induction configuration (e). Panels (f) and (g) show point-by-point matching profiles illustrating induced brightness/lightness structure in the test fields (open symbols) and the luminance profiles of the inducing grating (solid gray lines). The filled symbols in (f) show point-by-point brightness/lightness matches to the inducing grating. These matching profiles reveal that horizontal induced brightness/lightness gradient structure is present in the test fields of all of these stimuli (i.e., none of the test fields appear homogeneous). Predictions of the multiscale DOG filtering model of Blakeslee and McCourt (1997) are shown as blue lines in panels (k) through (o). The model accounts for both the structure and magnitude of the brightness/lightness variations observed in these stimuli. Figure belongs to M. E. McCourt and B. Blakeslee, and is based on data from Blakeslee & McCourt (1997).

accounted for by the output of a simple, multiscale difference of Gaussians (DOG) filtering model (Fig. III.55-1k–o). Using this same matching technique, Blakeslee and McCourt (1999) later showed that induced brightness/lightness was inhomogeneous in the test fields of White stimuli as well. The addition of orientation selectivity and response normalization to the DOG model, which resulted in the oriented DOG (ODOG) model, allowed the model to account for anisotropic effects such as the White effect as well (Fig. III.55-2).

That induction in the test fields of these stimuli is not homogeneous is of theoretical importance because virtually all current models of brightness/lightness perception, except the ODOG and other closely related models relying on spatial filtering (Blakeslee & McCourt, 1999; Robinson, Hammond, & de Sa, 2007), predict and assume that induction is homogeneous. They cannot, therefore, account for any of the induced brightness/lightness gradient effects. These models include both low-level filtering models based on the filling-in or assignment of brightness/lightness derived from information extracted by small receptive fields (i.e., "edge detectors"; Cornsweet & Teller, 1965; Fiorentini, Baumgartner, Magnussen, Schiller, & Thomas, 1990; Gerrits & Vendrik, 1970; Grossberg, 2003; Grossberg & Hong, 2006; Grossberg & Todorović, 1988; Kingdom, 2011; Kingdom & Moulden, 1988; Komatsu, 2006; Paradiso & Hahn, 1996; Paradiso & Nakayama, 1991; Rossi & Paradiso, 1996, 2003; Rossi, Rittenhouse, & Paradiso, 1996; Rudd, 2010; Rudd & Zemach, 2004), as well as models based on junction analysis (Adelson, 2000), Gestalt grouping (Gilchrist, 2006; Gilchrist et al., 1999), and perceptual inferences about depth, shape, illumination, reflectance, and transparency (Adelson, 2000; Anderson & Winawer, 2008; Bressan, 2006; Gilchrist, 2006; Gilchrist et al., 1999; Logvinenko & Ross, 2005).

Further challenges to brightness/lightness models based on a slow filling-in of information from the edges of the test field (Rossi & Paradiso, 1996, 2003) come from more recent studies of the temporal properties of grating induction. Whereas high-contrast sinewave luminance gratings have a distorted brightness/lightness profile, with narrow troughs and broad peaks, the brightness/lightness profile of induced gratings is veridically sinusoidal (McCourt, 1994). This property allowed Blakeslee and McCourt (2008) to use a quadrature phase motion paradigm to precisely measure the phase (time) lag of

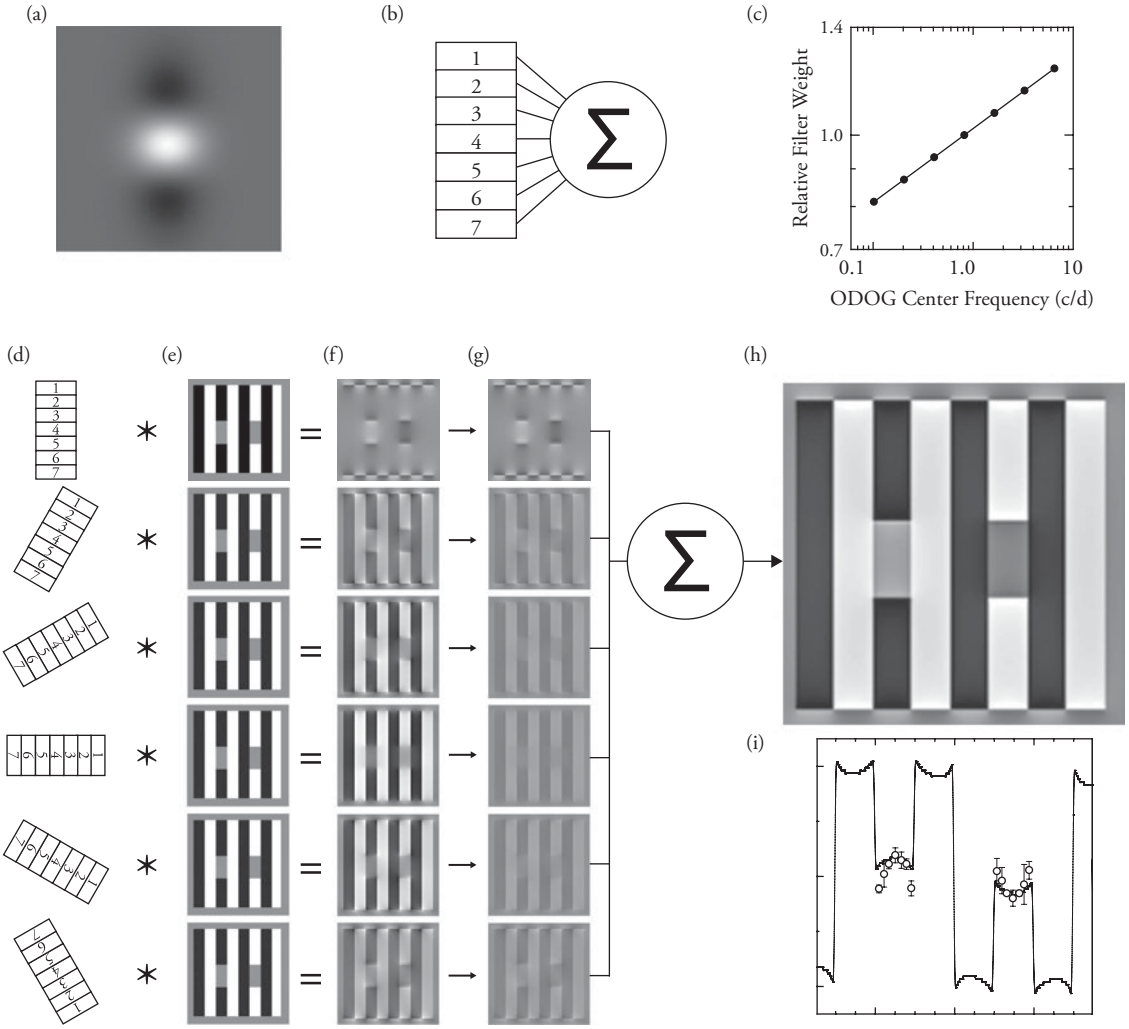


Figure III.55-2. ODOG model applied to White's effect. Panel (a) illustrates the spatial structure of an ODOG filter. Panel (b) indicates that seven filters at octave-interval spatial scales are summed at each filter orientation. Panel (c) shows the center frequencies of the seven ODOG filters and their relative weighting. Panel (d) illustrates the six channel orientations of the ODOG filters summed across scale. Panel (e) shows the White stimulus. Panel (f) shows the result of convolving each stimulus with the multiscale orientation channels shown in (d). Note the relatively higher level of activity for the horizontally oriented channel. Panel (g) shows root mean squared normalized activity in each orientation channel. Panel (h) shows the summed output across all six normalized orientation channels. Panel (i) plots a cross-section of ODOG model output to the White stimulus (black line) and point-by-point brightness/lightness matches (open symbols) across the two equiluminant test patches, which demonstrate spatial inhomogeneity. Figure belongs to M. E. McCourt and B. Blakeslee, and is based on data from Blakeslee & McCourt (1999).

grating induction at four temporal frequencies (2 Hz, 8 Hz, 16 Hz, and 24 Hz). They found that the temporal response of induced brightness/lightness differed from that of luminance gratings by a small temporal phase lag (0.016 cycle) or by a small time lag (1 ms), and remained constant across wide variations of temporal frequency and test field height (0.5°, 3.0°, 6.0°, and 9°). The findings that the phase (time) lag of grating induction did not vary with test field height and that induction was present even at a counterphase flicker rate of 24 Hz both argued strongly against any form of slow "fill-in" explanation for induction (DeValois, Webster, DeValois, & Lingelbach, 1986; Rossi & Paradiso, 1996). That this conclusion is not limited to grating induction but may apply to simultaneous contrast and the White effect as well is supported by recent studies (Kaneko & Murakami, 2012; Robinson & de Sa, 2008) that measured induction magnitude in these stimuli for very brief stimulus presentations.

Subsequently, Blakeslee and McCourt (2011) used the quadrature phase motion technique, in combination with a canceling technique, to characterize the entire *spatiotemporal surface* of grating induction by measuring its magnitude in dynamic displays across a wide range of spatial frequencies (0.03–4 c/d), temporal frequencies (0.25–16 Hz), and test field heights (0.5°–4°). In agreement with and extending previous results (Foley & McCourt, 1985; McCourt, 1982), canceling contrast was found to be a low-pass function of spatial and temporal frequency. Induction was greatest at low frequencies and declined with increasing frequency in both dimensions. Induction magnitude also declined with increasing test field height and was found to

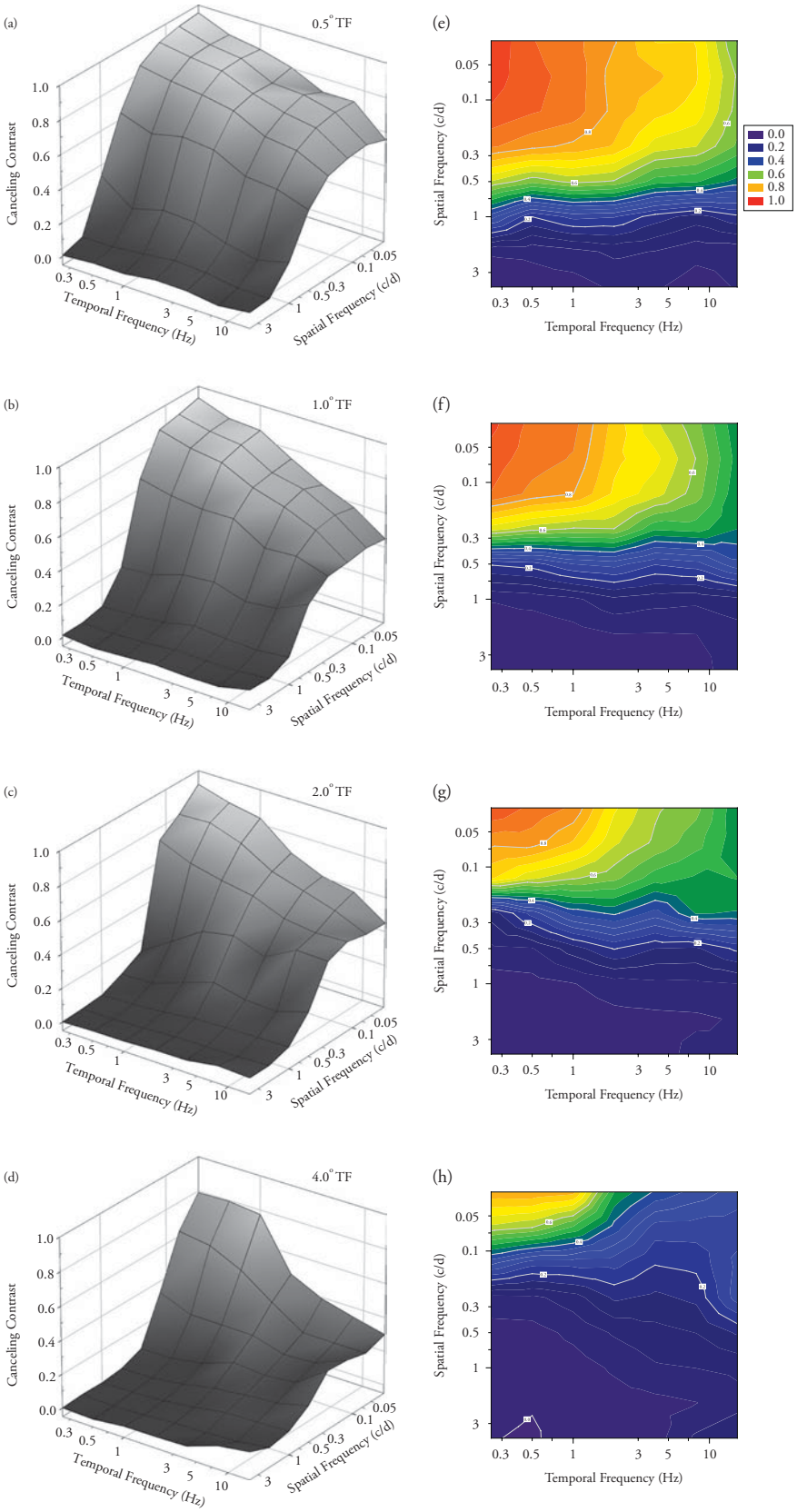


Figure III.55-3. Mesh (a)–(d) and contour plots (e)–(h) showing mean canceling contrast for four observers as a function of inducing grating spatial and temporal frequency. Color indexes canceling contrast in the contour plots. The rows plot data for the four test field heights 0.5°, 1.0°, 2.0°, and 4.0°, respectively. At all test field heights canceling contrast is a low-pass function of spatial and temporal frequency. Induction is greatest at low frequencies and declines with increasing frequency in both dimensions. Figure belongs to M. E. McCourt and B. Blakeslee, and is based on data from Blakeslee & McCourt (2011).

scale as the product of inducing grating spatial frequency (c/d) and test field height (Fig. III.55-3). Again, these properties of the data are readily accounted for by multiscale spatial filtering and lend further support to such models of brightness/lightness induction (Blakeslee & McCourt, 1997, 1999; Dakin & Bex, 2003; Foley & McCourt, 1985; McCourt, 1982; Robinson et al., 2007).

Most recently, in an effort to further distinguish the viability of the various competing models, Blakeslee and McCourt (2013) investigated the spatial characteristics of grating induction as a function of the distance from the inducing field edge using a narrow-probe quadrature-phase motion-cancellation paradigm and a point-by-point brightness/lightness matching paradigm. In particular, multiscale spatial filtering models such as the ODOG and related models (Blakeslee & McCourt, 1997, 1999; Robinson et al., 2007) and filling-in models (Grossberg & Todorović, 1988; Rossi & Paradiso, 2003) make very different predictions regarding the magnitude of induction as a function of this distance. Filling-in explanations predict that the brightness/lightness of the filled-in region will be homogeneous, whereas multiscale filtering predicts a fall-off of grating induction magnitude with distance from the inducing edge.

The spatial data obtained using both techniques (Figures III.55-4 and III.55-5) indicate that the magnitude of grating induction decreases with distance from the inducing edge for both 3° and 9° test field heights (Blakeslee & McCourt, 2013). For example, mean cancelling contrast decreased from a value of 0.35 at the inducing edge to an asymptotic value of approximately 0.07 over a distance of 1.5° (Fig. III.55-4). These results complement prior point-by-point brightness/lightness matching data for grating induction, simultaneous contrast, and White stimuli, which revealed the horizontal (as opposed to the vertical) induced structure in the test fields of these stimuli (Blakeslee & McCourt, 1997, 1999). In addition, the results confirm recent behavioral and functional magnetic resonance imaging evidence (Pereverzeva & Murray, 2008) that the magnitude of induction declines with distance from the inducing field edge of static central disks in which the disk surround was modulated at 1 Hz. The induced structure observed in all of these studies argues strongly against homogeneous filling-in mechanisms of any speed as well as against any mechanisms that posit that induction is homogeneous. The structured induction, however, is well accounted for by the ODOG model (Figures III.55-4 and III.55-5). Blakeslee and McCourt reported that the best fit predictions of the

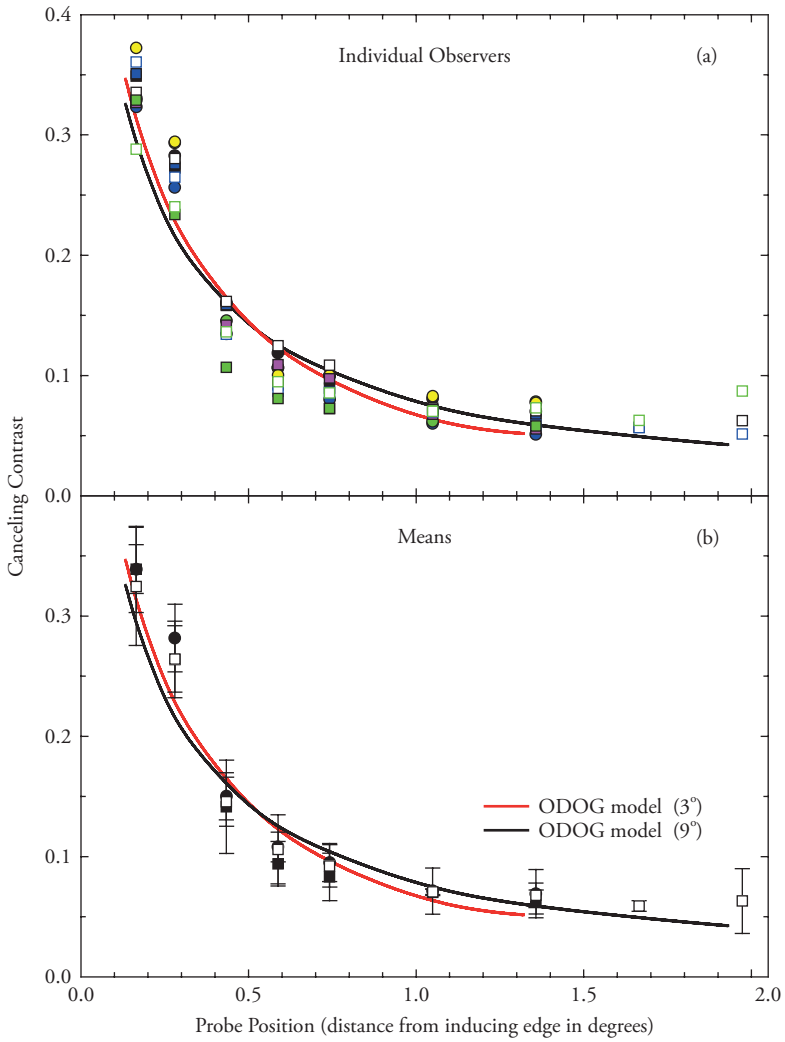


Figure III.55-4. Canceling contrast plotted as a function of distance from the inducing field edge measured using the quadrature-phase motion cancellation technique. The symbols in panel (a) plot the results from the four individual observers (indicated by color) for the 4 Hz, 3° test field height (filled circles); 1 Hz, 3° test field height (filled squares); and 1 Hz, 9° test field height (open squares) conditions. The symbols in panel (b) plot the means and 95% confidence intervals for these four observers for the same conditions: 4 Hz, test field 3° (filled circles); 1 Hz, test field 3° (filled squares); 1 Hz, test field 9° (open squares). The best fit predictions of the ODOG model (Blakeslee & McCourt, 1999) for the 3° and 9° conditions are indicated by the red and black lines, respectively, in both panels and account for 95% of the overall variance. Reprinted with permission from Blakeslee & McCourt (2013).

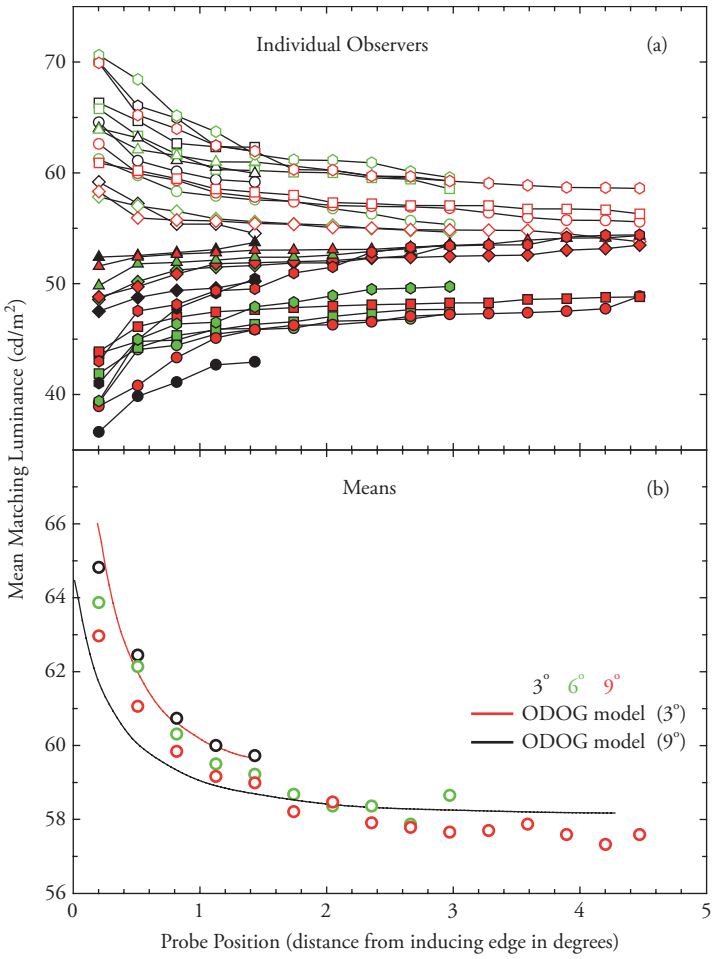


Figure III.55-5. Matching luminance as a function of distance from the inducing field edge is plotted separately for five observers (indicated by symbol shape) in panel (a). Matches were obtained for test field heights of 3° (black symbols), 6° (green symbols), and 9° (red symbols) at both the maximum (filled symbols) and minimum (open symbols) inducing grating phase. All subjects show a decrease in the magnitude of induction as a function of the distance from the inducing field edge, despite individual differences in overall induction magnitude. In panel (b) the data from the minimum inducing grating phase are collapsed across observers and plotted as means along with the linear best fit predictions of the ODOG model for the 3° and 9° conditions. The model predictions are a good fit to the data, accounting for 95% of the overall variance in the mean magnitude of induction found with distance from the inducing field edge. Reprinted with permission from Blakeslee & McCourt (2013).

ODOG model accounted for 95% of the overall variance in the mean magnitude of induction found with distance from the inducing field edge.

In light of these findings, it is reasonable to ask why induction is so often assumed to be homogeneous. Heinemann (1972) suggested that this may come about because observers can use different criteria when asked to assess the brightness of an extended area. In other words, the criterion for what an observer reports as homogeneous can be influenced by the question posed or the particular task given. By this reasoning, the grating induction configuration is a very effective one for revealing structured induction, at least in the horizontal direction, while the brightness gradients in the vertical direction of the grating induction stimulus and in other induction stimuli such as the White effect are more subtle, resulting in their description as homogeneous or structured depending on the question or task.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health Grants NIH R01 EY014015 (BB) and NIH P20 GM103505 (MEM). The National Institute of General Medical Sciences and the National Eye Institute are components of the National Institutes of Health. The contents of this report are solely the responsibility of the authors and do not necessarily reflect the official views of the National Institutes of Health, National Institute of General Medical Sciences, or National Eye Institute. Commercial relationships: None.

REFERENCES

Adelson, E. H. (2000). Lightness perception and lightness illusions. In M. Gazzaniga (Ed.), *The new cognitive neurosciences* (2d ed., pp. 339–351). Cambridge, MA: MIT Press.

Anderson, B. L., & Winawer, J. (2008). Layered image representations and the computation of surface lightness. *Journal of Vision*, *8*(7), 18, 1–22.

Blakeslee, B., & McCourt, M. E. (1997). Similar mechanisms underlie simultaneous brightness contrast and grating induction. *Vision Research*, *37*, 2849–2869.

Blakeslee, B., & McCourt, M. E. (1999). A multiscale spatial filtering account of the White effect, simultaneous brightness contrast and grating induction. *Vision Research*, *39*, 4361–4377.

Blakeslee, B., & McCourt, M. E. (2005). A multiscale filtering explanation of gradient induction and remote brightness induction effects: A reply to Logvinenko (2003). *Perception*, *34*(7), 793–802.

Blakeslee, B., & McCourt, M. E. (2008). Nearly instantaneous brightness induction. *Journal of Vision*, *8*(2), 15, 1–8, http://journalofvision.org/8/2/15/, doi:10.1167/8.2.15

Blakeslee, B., & McCourt, M. E. (2011). Spatiotemporal analysis of brightness induction. *Vision Research*, *51*, 1872–1879.

Blakeslee, B., & McCourt, M. E. (2013). Spatiotemporal analysis of brightness induction as a function of distance from the inducing field edge. *Vision Research*, *78*, 39–45.

Bressan, P. (2006). The place of white in a world of grays: A double-anchoring theory of lightness perception. *Psychological Review*, *113*, 526–553.

Chevreul, M. E. (1890). In C. Martel (Trans.), *The principles of harmony and contrast of colours*. London: Bell.

Cornsweet, T. N., & Teller, D. (1965). Relation of increment thresholds to brightness and luminance. *Journal of the Optical Society of America*, *55*, 1303–1308.

Dakin, S. C., & Bex, P. J. (2003). Natural image statistics mediate brightness "filling-in." *Proceedings of the Royal Society of London, B*, *270*, 2341–2348.

DeValois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, *22*, 545–559.

DeValois, R. L., & DeValois, K. K. (1988). *Spatial vision*. New York: Oxford University Press.

DeValois, R. L., Webster, M. A., DeValois, K. K., & Lingelbach, B. (1986). Temporal properties of brightness and color induction. *Vision Research*, *26*, 887–897.

Fiorentini, A., Baumgartner, G., Magnussen, S., Schiller, P. H., & Thomas, J. P. (1990). The perception of brightness and darkness: Relations to neuronal receptive fields. In L. Spillmann & J. S. Werner (Eds.), *Visual perception: The neurophysiological foundations* (pp. 129–161). San Diego, CA: Academic Press.

Foley, J. M., & McCourt, M. E. (1985). Visual grating induction. *Journal of the Optical Society of America A*, *2*, 1220–1230.

Geier, J., & Hudak, M. (2011). Changing the Chevreul illusion by a background luminance ramp: Lateral inhibition fails at its traditional stronghold—A psychophysical refutation. *PLoS ONE*, *6*(10), e26062. doi:10.1371/journal.pone.0026062

Geisler, W. S., Albrecht D. G., & Crane A. M. (2007). Responses of cortical neurons to transient changes in local contrast and luminance. *Journal of Neuroscience*, *27*, 5063–5067.

Gerrits, H. J. M., & Vendrik, A. J. H. (1970). Simultaneous contrast, filling-in process and information processing in man's visual system. *Experimental Brain Research*, *11*, 411–430.

Gilchrist, A. L. (2006). *Seeing black and white*. New York: Oxford University Press.

Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., . . . Economou, E. (1999). An anchoring theory of lightness perception. *Psychological Review, 106*, 795–834.

Grossberg, S. (2003). Filling-in the forms: Surface and boundary interactions in visual cortex. In L. Pessoa & P. DeWeerd (Eds.), *Filling-in: From perceptual completion to cortical reorganization* (pp. 13–37). New York: Oxford University Press.

Grossberg, S., & Hong, S. (2006). A neural model of surface perception: Lightness, anchoring, and filling-in. *Spatial Vision*, *19*(2–4), 263–321.

Grossberg, S., & Todorović, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena. *Perception & Psychophysics*, *43*, 241–277.

Heinemann, E. G. (1972). Simultaneous brightness induction. In D. H. Jameson & L. M. Hurvich (Eds.), *Handbook of sensory physiology* (Vol. VII/4, pp. 147–169). Berlin: Springer.

Hermann, L. (1870). Eine erscheinung simultanen contrastes. *Pflügers Archiv für die gesamte Physiologie*, *3*, 13–15.

Kingdom, F. A. A. (1999). Old wine in new bottles? Some thoughts on Logvinenko's "Lightness induction revisited." *Perception*, *28*, 929–934.

Kingdom, F. A. A. (2011). Lightness, brightness and transparency: A quarter century of new ideas, captivating demonstrations and unrelenting controversy. *Vision Research*, *51*, 652–673.

Kingdom, F. A. A., & Moulden, B. (1988). Border effects on brightness: A review of findings, models and issues. *Spatial Vision*, *3*, 225–262.

Kingdom, F., & Moulden, B. (1992). A multi-channel approach to brightness coding. *Vision Research*, *32*, 1565–1582.

Komatsu, H. (2006). The neural mechanisms of perceptual filling-in. *Nature Reviews Neuroscience*, *7*, 220–231.

Logvinenko, A. D., & Ross, D. (2005). Adelson's tile and snake illusions: A Helmholtzian type of simultaneous lightness contrast. *Spatial Vision*, *18*(1), 25–72.

McCourt, M. E. (1982). A spatial frequency dependent grating-induction effect. *Vision Research*, *22*, 119–134.

McCourt, M. (1994). Grating induction: A new explanation for stationary visual phantoms. *Vision Research*, *34*, 1609–1618.

McCourt, M. E., & Blakeslee, B. (1994). A contrast matching analysis of grating induction and suprathreshold contrast perception. *Journal of the Optical Society of America A*, *11*, 14–24.

McCourt, M. E., & Kingdom, F. A. A. (1996). Facilitation of luminance grating detection by induced gratings. *Vision Research*, *36*, 2563–2574.

Moulden, B., & Kingdom, F. A. A. (1989). White's effect: A dual mechanism. *Vision Research*, *29*, 1245–1259.

Moulden, B., & Kingdom, F. A. A. (1991). The local border mechanism in grating induction. *Vision Research*, *31*, 1999–2008.

Paradiso, M. A., & Hahn, S. (1996). Filling-in percepts produced by luminance modulation. *Vision Research*, *36*, 2657–2663.

Paradiso, M. A., & Nakayama, K. (1991). Brightness perception and filling-in. *Vision Research*, *31*, 1221–1236.

Ratliff, F. (1965). *Mach bands: Quantitative studies on neural networks in the retina* Holden-Day Series in Psychology. San Francisco, CA: Holden-Day.

Robinson, A. E., & de Sa, V. R. (2008). Brief presentations reveal the temporal dynamics of brightness induction and White's Illusion. *Vision Research*, *48*, 2370–2381.

Robinson, A. E., Hammond, P. S., & de Sa, V. R. (2007). Explaining brightness illusions using spatial filtering and local response normalization. *Vision Research*, *47*, 1631–1644.

Rossi, A. F., & Paradiso, M. A. (1996). Temporal limits of brightness induction and mechanisms of brightness perception. *Vision Research*, *36*, 1391–1398.

Rossi, A. F., & Paradiso, M. A. (2003). Surface completion: Psychophysical and neurophysiological studies of brightness. In L. Pessoa & P. DeWeerd (Eds.), *Filling-in: From perceptual completion to cortical reorganization* (pp. 59–105). New York: Oxford University Press.

Rossi, A. F., Rittenhouse, C. D., & Paradiso, M. A. (1996). The representation of brightness in primary visual cortex. *Science*, *273*, 1104–1107.

Rudd, M. E. (2010). How attention and contrast gain control interact to regulate lightness contrast and assimilation: A computational neural model. *Journal of Vision*, *10*(14), 40, 1–37. doi:10.1167/10.14.40.

Rudd, M. E., & Zemach, I. K. (2004). Quantitative properties of achromatic color induction: An edge integration analysis. *Vision Research*, *44*, 971–981.

Shapley, R., & Reid, R. C. (1985). Contrast and assimilation in the perception of brightness. *Proceedings of the National Academy of the Sciences USA*, *82*, 5983–5986.

Spillmann, L. (1994). The Hermann grid illusion: A tool for studying human perceptive field organization. *Perception*, *23*, 691–708.

Wilson, H. R., & Wilkinson, F. (2003). Spatial channels in vision and spatial pooling. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neurosciences* (pp. 1060–1068). Cambridge, MA: MIT Press.

Zaidi, Q. (1989). Local and distal factors in visual grating induction. *Vision Research*, *29*, 691–697.

Chapter 56
Illusory Variations in Apparent Dot Brightness Induced by Density Modulations

Jeffrey B. Mulligan

This illusion demonstrates that we are unable to judge the brightness of a small dot independently of the total amount of light falling in its local neighborhood. This can be seen in Figure III.56-1: all of the dots have been rendered with the same intensity, but the dots appear brighter in the regions where the density is higher. Figure III.56-1 depicts a "plaid" of density modulation applied to a regular lattice of dots. The position of each dot is perturbed horizontally and vertically according to a pair of sinusoidal functions. The amplitude of the modulation decreases gradually, moving up and to the right from the lower left-hand corner, by a factor of 2 for each cycle of the modulation.

I first observed this effect in 1982 during an extended visit to AT&T Bell Laboratories in Murray Hill, New Jersey. Using equipment in the laboratory of Bela Julesz, an attempt was made to measure minimum motion thresholds for sinusoidal modulation of compressive motion, analogous to the thresholds for shearing motion previously reported by Nakayama and Tyler (1981). A shear grating is generated by modulating a velocity in the perpendicular direction, while in a compression grating the modulation is in the same direction. Examples of these types of modulation abound in nature. For example, water waves are an example of transverse waves, analogous to a shear grating: the motion of individual bits of water is up and down, while the height of the surface varies in the horizontal dimension. Sound waves, on the other hand, are an example of a longitudinal wave, analogous to a compression grating. In a sound wave, the particles of the medium move back and forth in the same direction as the modulation, producing alternating zones of compression and rarifaction. Taken together, sinusoidal velocity gratings of shear and compression form a complete Fourier basis for velocity fields. Unlike shear gratings, compression gratings do not preserve the local density of the underlying texture; as the name implies, the texture undergoes local compression and expansion.

When the first compression grating was created, a surprising observation was made: under certain conditions, small movements of the dots could be detected, not via the perception of movement of the elements but rather via a change in apparent brightness induced by the motion, even when the motion itself was invisible. Some observers may see a static analog of this in Figure III.56-1: in the region with the smallest perceptible modulation, is the modulation seen as a change in brightness, a change in dot spacing, or both? In the case where a threshold motion is detected via the illusory brightness change, the modulation can be rendered invisible by applying a compensatory luminance modulation that maintains a constant space-average luminance. When the modulation amplitude of both the motion and the luminance compensation is increased, it eventually becomes visible as motion of dots with constant apparent brightness (Nakayama, Silverman, MacLeod, & Mulligan, 1985).

The perceived brightness can be made to appear constant if the space-average luminance is made constant. This can be done in two distinct ways, either by applying a compensatory modulation to the luminances of the dots (lowering the luminances of the dots that are closer together; see Figure III.56-2) or applying a background modulation that adds just the right amount of luminance to the regions with fewer dots (Fig. III.56-3).[1] The compensating background used in Figure III.56-3 is shown in isolation without the dots in Figure III.56-4. This background was computed using an iterative procedure: the average luminance of the dot array was first computed by blurring. This image was then inverted and adjusted to have a minimum value of zero. A new image was then formed by taking the maximum value from the original dot array and the computed background. Because the background values are discarded at the dot locations, the compensation will not be perfect, but the computation can be repeated, with the new, smaller compensation added to the previous background. This process is repeated until the corrections are insignificant.

Empirical measurements documenting the effect can be found in Mulligan and MacLeod (1988). Here we summarize their findings: constant perceived brightness can be obtained under reciprocal modulations of density and luminance for average dot spacings of 7.5 and 15 min of arc, but when the spacing is 30 min or more, the brightness is unaffected by density modulations and no compensating luminance modulation is required to achieve perceptually

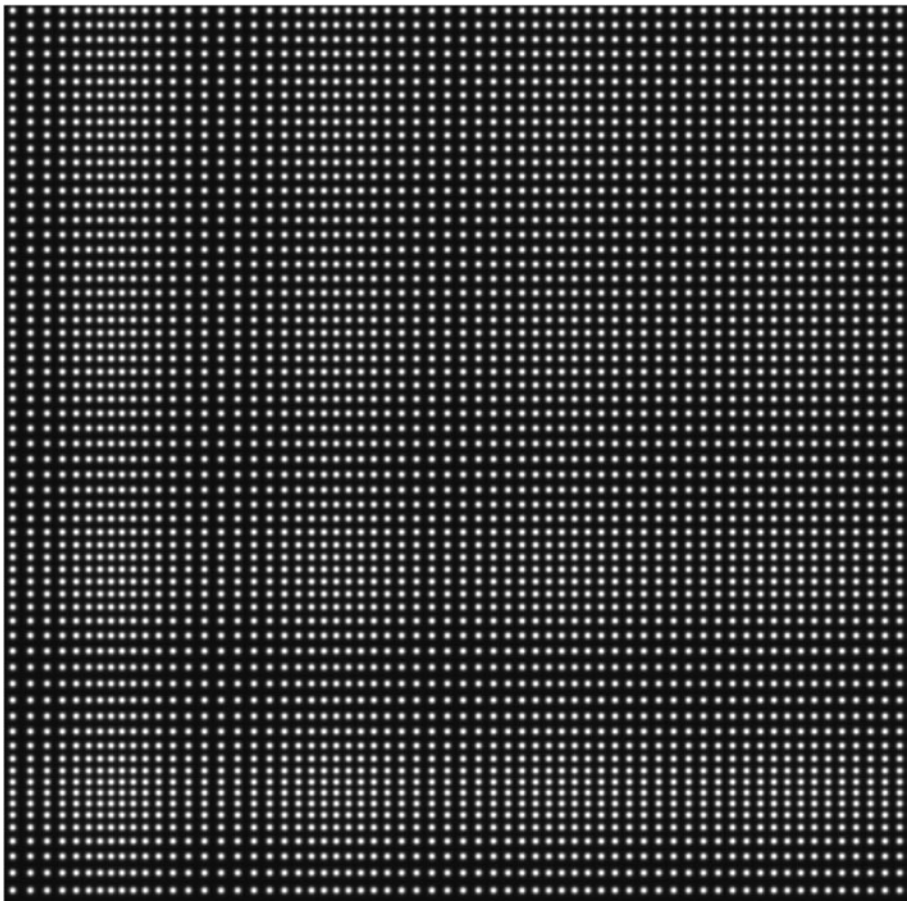


Figure III.56-1. Density modulation of constant intensity dots produces an illusory increase in dot brightness in regions of high density. Image by Jeffrey B. Mulligan.

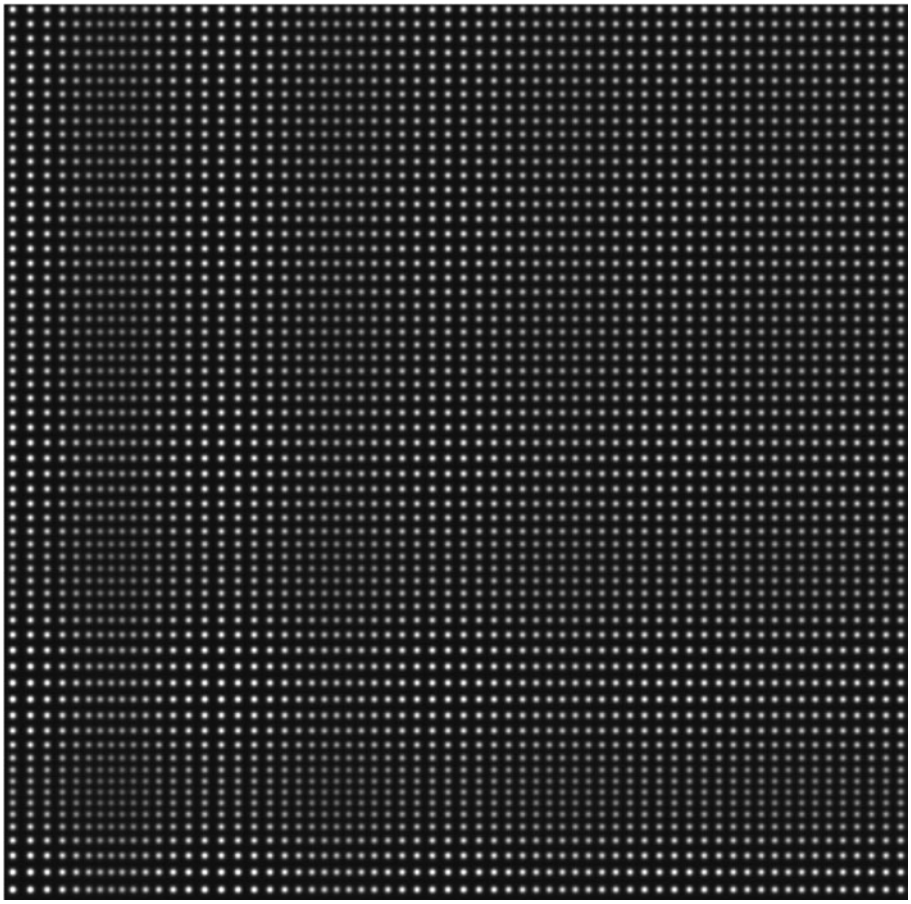


Figure III.56-2. Cancellation of the effect by reciprocal modulation of dot intensity. (The precise gray levels needed to demonstrate cancellation may not be reproduced correctly.) Image by Jeffrey B. Mulligan.

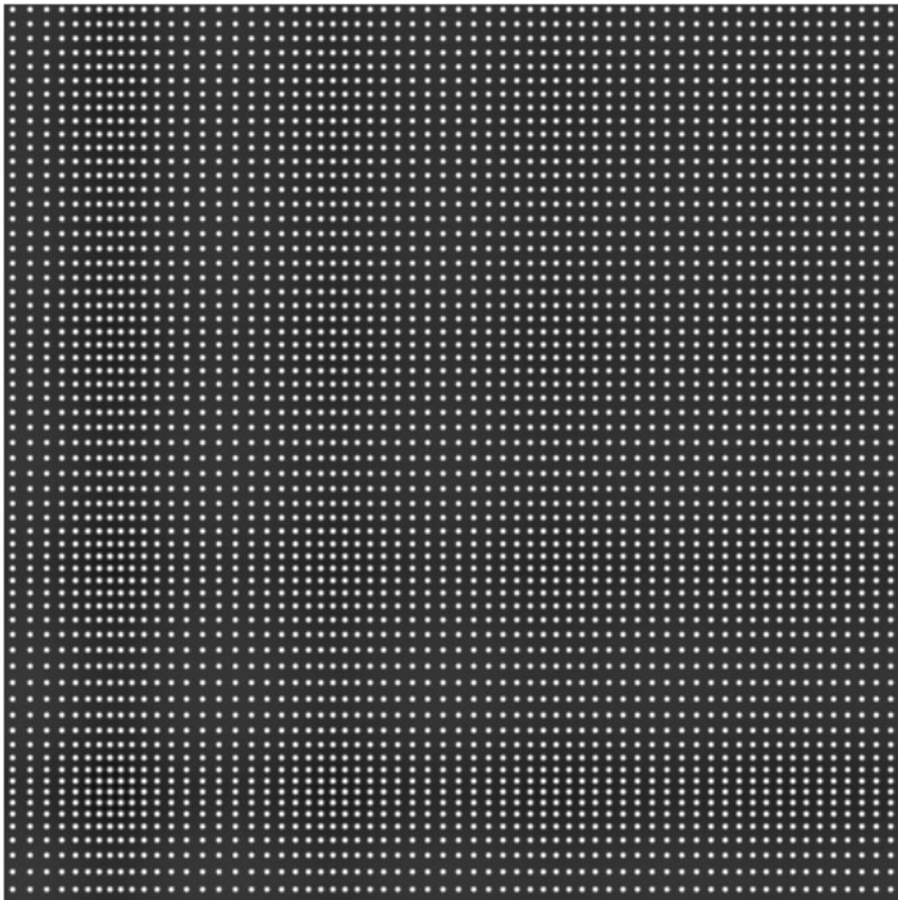


Figure III.56-3. Cancellation of the effect by incorporation of a compensating background. (The precise gray levels needed to demonstrate cancellation may not be reproduced correctly.) Image by Jeffrey B. Mulligan.



Figure III.56-4. The compensating background used in Figure III.56-3. Image by Jeffrey B. Mulligan.

uniform brightness (Experiment 1); detection thresholds for density modulation are increased in the presence of reciprocal luminance modulation and decreased for in-phase luminance modulation (Experiments 2 and 3); detection thresholds for luminance modulation applied to a regular array of dots are elevated in the presence of a compensating background, and decreased for in-phase background modulation (Experiment 4).

This effect is interesting because it demonstrates that the attributes of a stimulus are not all represented at the same resolution as the contours of objects. The dots are all clearly resolved, and yet the brightness seems be drawn from an area (roughly half a degree of visual angle in diameter) that is much larger than an individual dot. It cannot be explained by a simple blurring of the image, as the dots appear sharp and the background does not seem to be affected. It is especially curious that the increased brightness that results from an increase in density is confined to the dots; while the effect can be cancelled by a modulation of either the dots themselves or the background, the effect is never attributed to the background.

To explain the effect, it is necessary to postulate that object properties such as brightness are represented independently from the contours that segment objects. Models with this property have been a recurrent theme in the work of Grossberg, as typified by a model proposed to explain effects such as subjective contours (e.g., Grossberg, Mingolla, & Todorović, 1989). To explain the present effect, however, the units conveying feature properties (such as brightness) must necessarily come from lower frequency channels, with receptive field sizes on the order of half of a degree.

NOTE

1. Complete cancellation of the illusion, by either method, requires precise control of display gray levels. As it was not possible to calibrate the printing process prior to publication of this chapter, the figures may not faithfully reproduce complete cancellation of the illusion.

REFERENCES

Grossberg, S., Mingolla, E., & Todorović, D. (1989). A neural network architecture for preattentive vision. *IEEE Transactions on Biomedical Engineering*, *36*, 65–84.

Mulligan, J. B., & MacLeod, D. I. A. (1988). Reciprocity between luminance and dot density in the perception of brightness. *Vision Research*, *28*, 503–519.

Nakayama, K., Silverman, G. H., MacLeod, D. I. A., & Mulligan, J. (1985). Sensitivity to shearing and compressive motion in random dots. *Perception*, *14*, 225–238.

Nakayama, K., & Tyler, C. W. (1981). Psychophysical isolation of movement sensitivity by removal of familiar position cues. *Vision Research*, *21*, 427–433.

Chapter 57
On the Watercolor Illusion

Baingio Pinna

THE WATERCOLOR ILLUSION

The watercolor illusion (Broerse & O'Shea, 1995; Broerse, Vladusich, & O'Shea, 1999; Pinna, 1988, 2005, 2006, 2008a, 2008b; Pinna, Brelstaff, & Spillmann, 2001; Pinna & Grossberg, 2005; Pinna & Tanca, 2008; Pinna, Werner, & Spillmann, 2003; Tanca & Pinna, 2008; Werner, Pinna, & Spillmann, 2007) shows two main phenomena: (a) a clear coloration effect (i.e., a long-range assimilative spread of color, originating from a thin chromatic contour juxtaposed to a darker colored contour) and (b) a strong figural effect, (i.e., figure–ground and object-hole segregations) across a large region.

In Figure III.57-1a, orange undulated contours juxtaposed to purple ones clearly appear as two unknown irregular shapes evenly colored by a light veil of orange tint (coloration effect). They are also perceived as segregated from the white background with a solid volumetric attribute, that is, as surfaces segregated in depth and extending out from the flat surface (figural effect). On the contrary, the complementary regions appear as empty spaces or holes (Pinna & Reeves, 2006).

In Figure III.57-1b, by reversing the orange and purple undulated contours of Figure III.57-1a, the Mediterranean seacoasts (a) and the Gulf of Mexico (b) emerge at once as figures, that is, the two seacoasts belong unilaterally to the surrounding main lands, both illusorily, filled with the

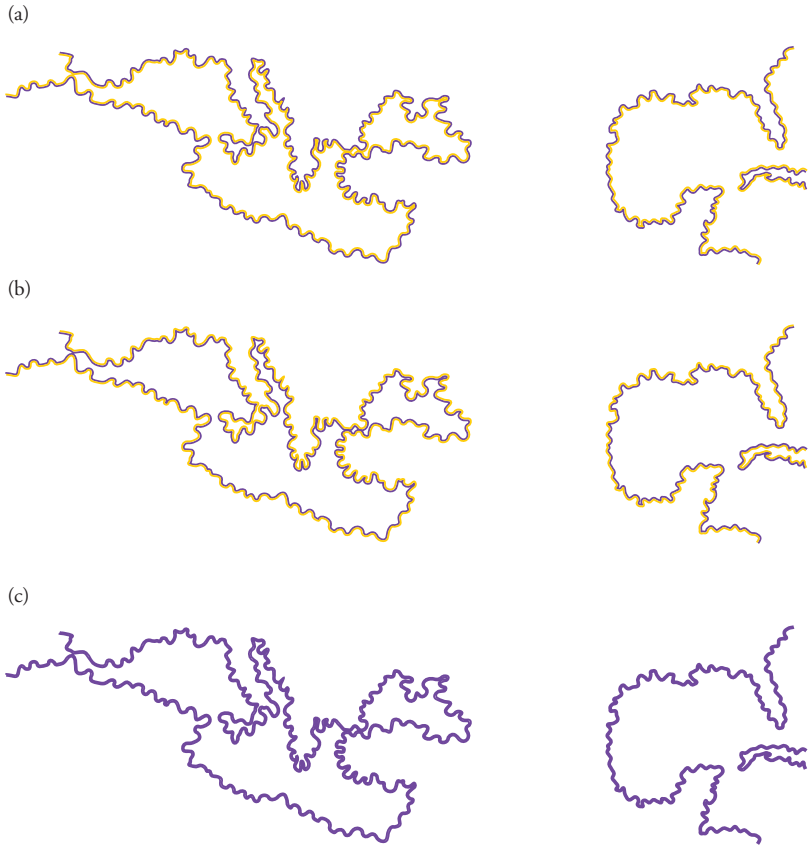


Figure III.57-1. Examples of the Watercolor illusion. (a-b) Orange contours juxtaposed to purple ones spread a light veil of orange tint (coloration effect) and induce a solid figure-ground and volumetric attribute (figural effect) to the colored region. Due to the watercolor illusion, the Mediterranean seacoasts and the Gulf of Mexico, invisible in (a), strongly emerge in (b). (c) A control. (Figures by Dr. Baingio Pinna's creations.)

same orange light veil as in Figure III.57-1a. What is figure and "visible" in Figure III.57-1b, is background and "invisible" in Figure III.57-1a. As in Figure III.57-1a, the two geographical main lands manifest a volumetric and a three-dimensional (3D) appearance (Pinna & Reeves, 2006).

In Figure III.57-1c, by replacing the orange contours with purple ones, the Mediterranean seacoasts (a) and the Gulf of Mexico (b) are perceived less clearly in terms of land versus sea (figure vs. background) than in Figure III.57-1b. The figure–ground segregation can be easily reversed and switched to the two unknown irregular shapes. The volumetric and coloration effects are absent (Pinna & Reeves, 2006).

The Coloration Effect

It was demonstrated (Pinna, 2008a, 2008b; Pinna et al., 2001) that the coloration effect shows the following main properties (Devinck, Delahunt, Hardy, Spillmann, & Werner, 2005, 2006; Devinck, Hardy, Delahunt, Spillmann, & Werner, 2006; Devinck & Knoblauch, 2012; Devinck & Spillmann, 2009; Devinck, Spillmann, & Werner, 2006; Pinna, 1988, 2005, 2006, 2008a, 2008b; Pinna et al., 2001; Pinna & Grossberg, 2005; Pinna & Tanca, 2008; Pinna et al., 2003; Tanca & Pinna, 2008; Werner et al., 2007).

1. It is approximately uniform.
2. It extends up to about 45° visual angle. Figure III.57-2a demonstrates that the perceived orange coloration fills evenly and strongly the large rectangle within a large range of observation distances and also by zooming in very closely.
3. The coloration effect fills in about 100 ms.
4. All the chromatic and achromatic colors in every possible combination can generate the effect, as shown in Figure III.57-2b, where the same irregular shape of Figure III.57-2a appears now filled with a light-green tint, and in Figure III.57-2c, where black and white juxtaposed contours induce a spectrum of gray tones.
5. The illusory coloration occurs on colored and black backgrounds as demonstrated in Figures III.57-2d and III.57-2e. In both conditions the orange tones (Fig. III.57-2d) and purple tones (Fig. III.57-2e), the colorations fill and mix, respectively, with the cyan and black backgrounds.
6. The optimal contour thickness is approximately 6 arc min.
7. Even if the coloration effect is stronger with irregular contours, it also occurs with straight contours and with chains of dots (cf. Figures III.57-2f and III.57-2g).
8. The contour with less luminance contrast relative to the background spreads proportionally more than the contour with a higher luminance contrast.
9. The color spreads in a direction orthogonal to the contour orientation.
10. By reversing the colors of the two juxtaposed contours, the coloration reverses accordingly (compare Figures III.57-1a and III.57-1b).
11. Although the direction of the color spreading due to both contours goes in opposite directions, the countours produce a coloration effect obtained by a combination of both colors in terms of saturation.
12. High luminance contrast between inducing contours shows the strongest coloration effect; however, it is clearly visible at near equiluminance as demonstrated in Figure III.57-2h.
13. Phenomenally, the coloration appears solid, impenetrable, and epiphanous as a surface color, as shown in all the previous figures. However, as illustrated in Figure III.57-3a, it can also appear transparent under conditions that are known to elicit transparent figures. A photographic watercolor version is shown in Figure III.57-3b, where a fisherman appears to be casting a fishing net that appears yellowy and transparent. In Figure III.57-3c, by replacing the shape of the watercolored net with an orange contour, the inner yellowy coloration and the soft realistic transparency of Figure III.57-3b are annulled. The coloration can also appear like a diaphanous fog, as demonstrated in Figure III.57-3d.
14. Similarly to neon color spreading (Pinna & Grossberg, 2005), the watercolor illusion induces a complementary color when one of the two juxtaposed contours is achromatic and the other chromatic, as shown in Figure III.57-3e.
15. The differences between the watercolor illusion and neon color spreading (Fig. III.57-3f) can be reduced under modified geometrical conditions (i, ii, iii, iv), and, by reaching a limiting case (iv), they can be eliminated. Figure III.57-3f shows conditions that introduce a limiting case (Pinna & Grossberg, 2005). If this is true, then while the similarities may be attributed to the local nearby transition of colors that are common to both illusions, the differences may be attributed to the global geometrical boundary conditions that differ in the two illusions, notably, the continuation of a segment in a different color in the neon color spreading and the juxtaposition of at least two lines in the watercolor illusion.

The Figure–Ground Effect

The figure–ground effect manifests the following properties (Pinna, 1988, 2005, 2006, 2008a, 2008b; Pinna et al., 2001; Pinna & Grossberg, 2005; Pinna & Tanca, 2008; Pinna et al., 2003; Tanca & Pinna, 2008; Werner et al., 2007).

1. The watercolor illusion strongly enhances the "unilateral belongingness of the boundaries," that is, a phenomenal property, first reported by Rubin (1921), according to which the boundaries belong only to the figure and not to the background, which appears as an empty space without a shape.
2. It follows that the figure–ground segregation induced by the watercolor illusion is not reversible and unequivocal like in the classical conditions studied by Gestalt psychologists. This is clearly demonstrated by comparing the two conditions as illustrated in Figure III.57-4. In Figure III.57-4a, the irregular purple contour is alternately and reversibly perceived like two peninsulas going from left to right with the mainland at the bottom or, alternatively, like two peninsulas going from right to left with the mainland at the top (Pinna & Tanca, 2008). Under watercolor

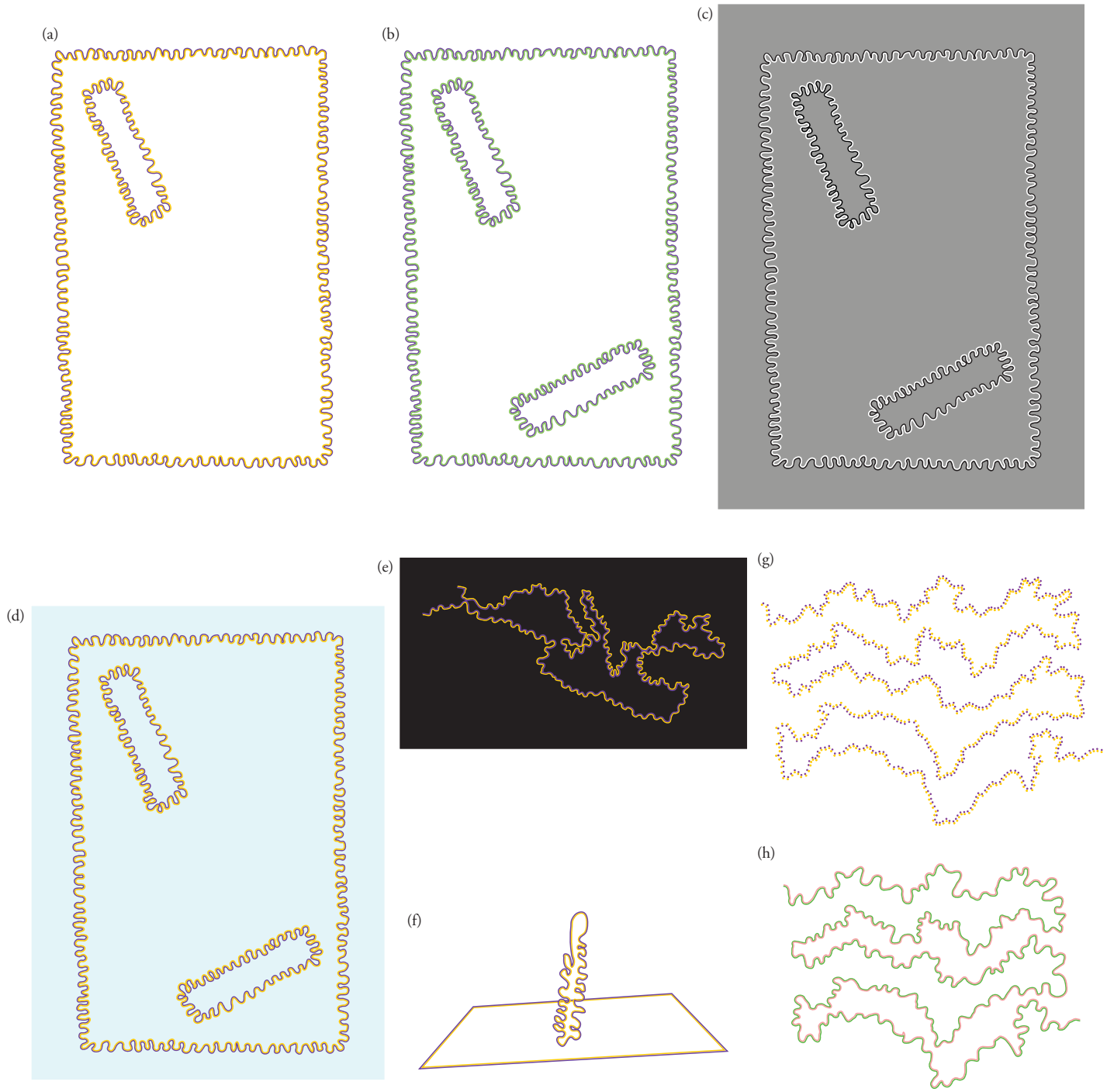


Figure III.57-2. The coloration effect of the Watercolor illusion. (a) Long range and uniform spreading up to about 45 deg visual angle. (b, c) All the chromatic and achromatic colors in every possible combination can generate the effect. (d, e) The illusory coloration occurs on colored and black backgrounds. (f-h) The coloration effect occurs with straight contours, with chains of dots and at near equiluminance. (Figures by Dr. Baingio Pinna's creations.)

conditions (Fig. III.57-4b), the irregular contours are now perceived much more strongly or uniquely as two peninsulas going from right to left connected to the mainland at the top.

3. Under watercolor conditions, both the figure and the background (cf. Figures III.57-4a and III.57-4b) assume new properties, becoming, respectively, bulging objects and empty spaces, both with a 3D volumetric appearance.
4. By reversing the colors of the two juxtaposed contours, the figure–ground segregation reverses accordingly. In Figure III.57-4c, the two complementary peninsulas of Figure III.57-4b going from left to right and connected to the mainland at the bottom now emerge. Figures III.57-4b and III.57-4c do not appear as the same figure with reversed chromatic edges but as two different figures. One possible result disappears inside the other due to the unilateral belongingness

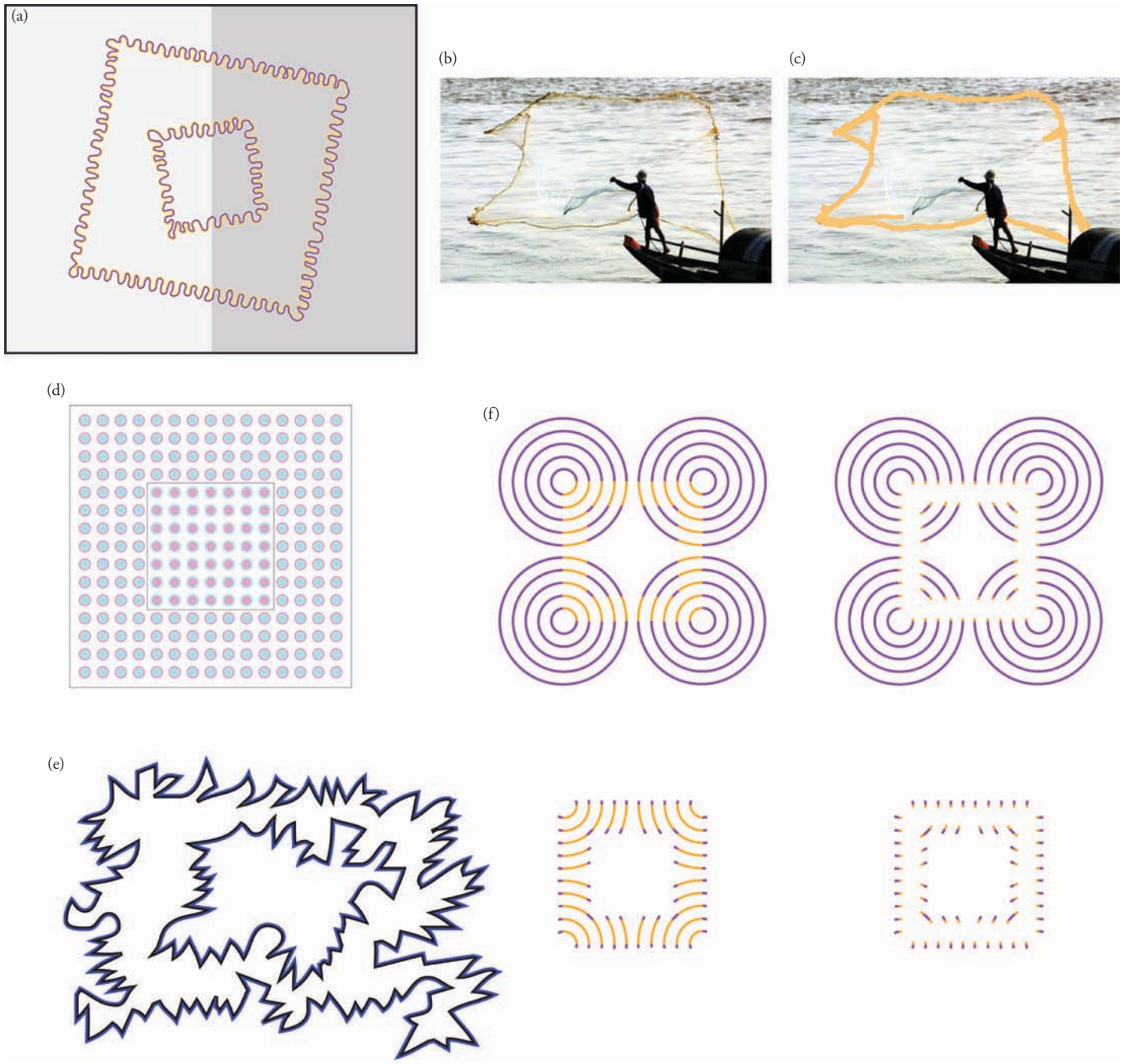


Figure III.57-3. The coloration effect of the Watercolor illusion. (a, b) The illusory coloration appears transparent. (c) A control. (d) The coloration appears like a diaphanous fog. (e) The watercolor illusion induces a complementary color: the inner region appears yellowish. (f) The differences between the watercolor illusion and the neon color spreading can be reduced to a common limiting case. (Figures by Dr. Baingio Pinna's creations.)

of the boundaries that in the watercolor illusion is strongly enhanced and placed in the 3D space.

In Videos III.57-1 and III.57-2, animated continuous variations of the color and the luminance contrast of the two juxtaposed contours show a continuous bulging as a figure of one region and the disappearing empty space of the complementary region and vice versa.

5. As shown in Figure III.57-4d, by gradually reversing the luminance contrast of the background (e.g., from white to black) while the color and luminance contrast of the juxtaposed contours is kept constant, the figure–ground segregation reverses: the orangish crosses at the top gradually become purplish stars at the bottom.
6. The watercolor illusion determines grouping and figure–ground segregation more strongly than the classical Gestalt principles. The principles of grouping by Wertheimer (1923) determine how elements group in parts that in their turn group in overall objects and how individual elements create larger wholes separated from others. The principles of figure–ground

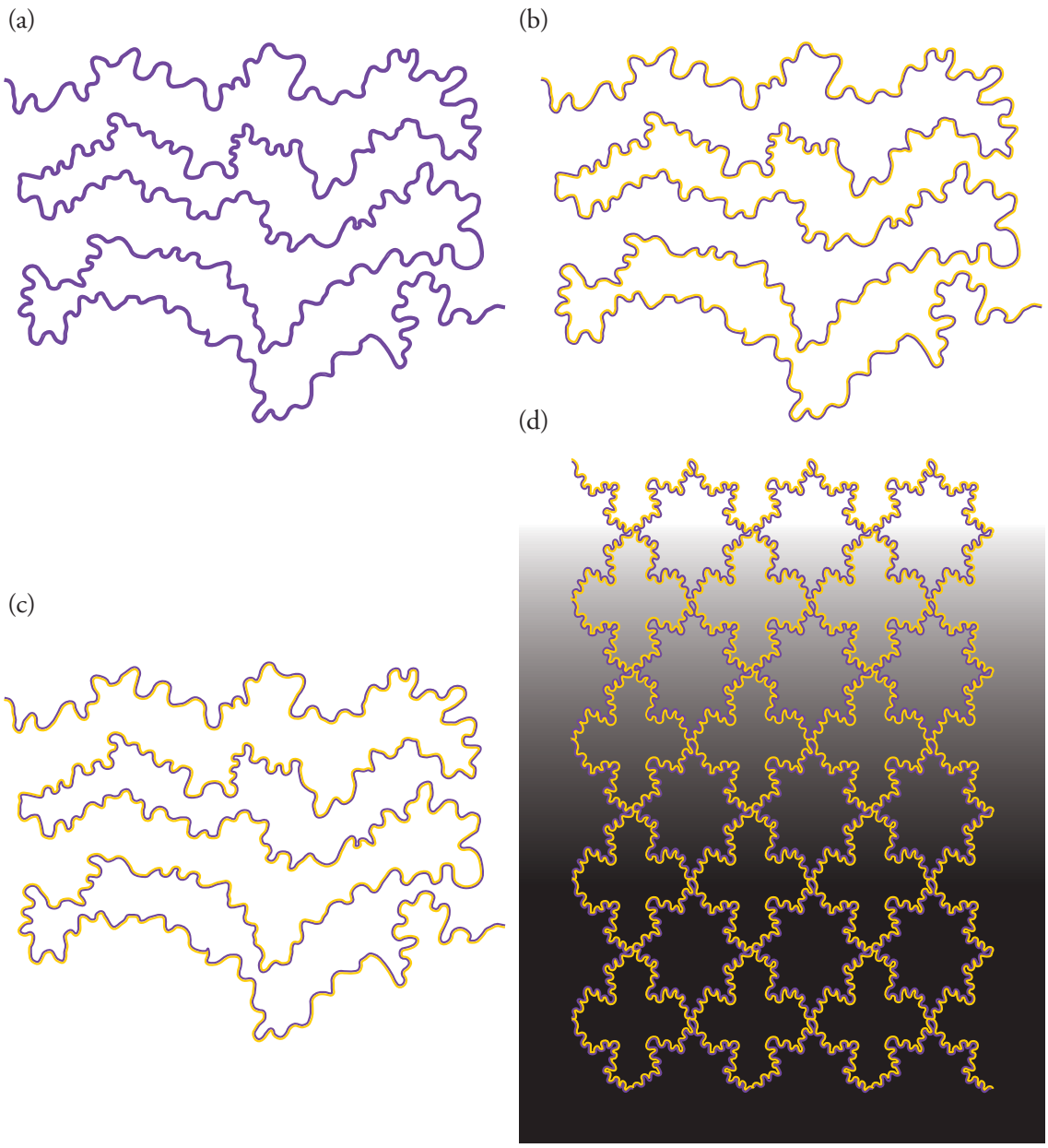


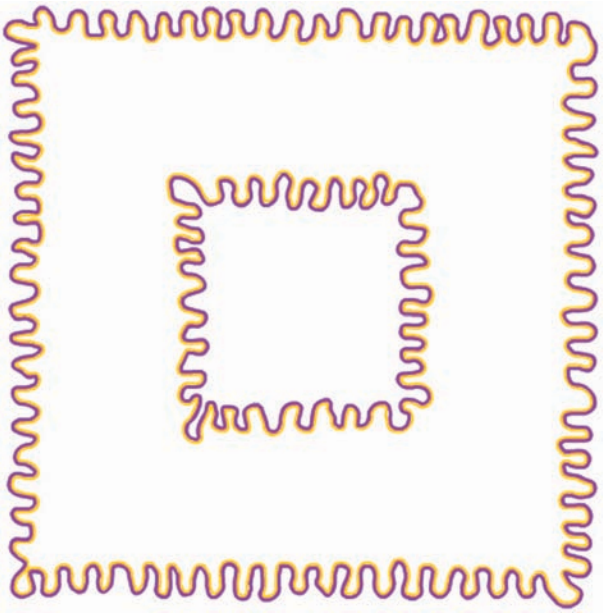
Figure III.57-4. The figure-ground effect of the Watercolor illusion. (a–c) The figure-ground segregation induced by the watercolor illusion is not reversible and unequivocal like in the classical conditions studied by Gestalt psychologists. (d) By gradually reversing the luminance contrast of the background, while both the color and luminance contrast of the juxtaposed contours is kept constant, the figure-ground segregation reverses and the orangish crosses at the top gradually become purplish stars at the bottom. (Figures by Dr. Baingio Pinna's creations).

segregation by Rubin (1921) represent the rules of belongingness of the boundaries to the figure and not to the ground, that is, what appears as a figure and what as a background. The watercolor illusion wins against all grouping and figure–ground principles as demonstrated in some of the next examples.

In Figure III.57-5a (left), according to the surroundedness figure–ground principle (some role is also played by the proximity grouping principle), the inner undulated square tends to appear as a figure segregated on a large background with a square-like shape. In Figure III.57-5a (right), the watercolor illusion reverses the previous figure–ground organization: the inner square is now perceived as a hole, and the region in between the two irregular squares appears as a figural solid frame.

In Figure III.57-5b (left), the role of the relative orientation and surroundedness principles is enhanced by the watercolor illusion; therefore the inner shape is more easily perceived as a figure on a larger empty background. The frame can very faintly be perceived as a figure. However, by reversing in Figure III.57-5b (right) the two chromatic adjacent contours of the watercolor illusion, the previous organization is reversed and the frame now pops out as a figure, while the inner space surrounded by the small irregular square appears as a hole.

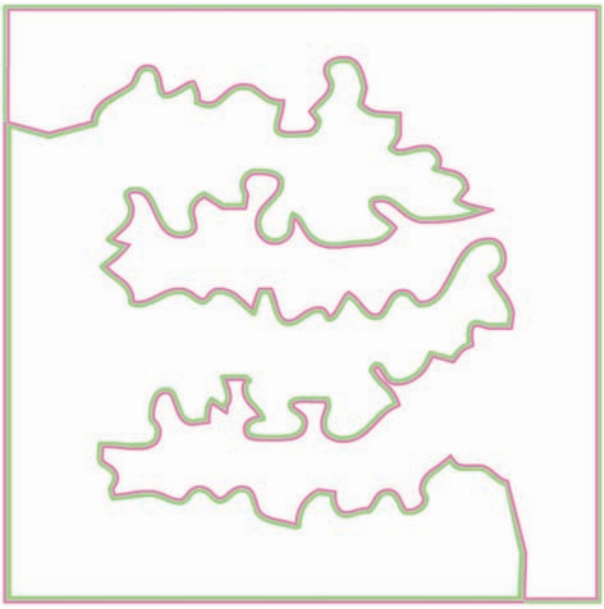
Figure III.57-5c shows the strength of the watercolor illusion in reversing the role of the regularity and Prägnanz.



Video III.57-1. Video by Dr. Bainglio Pinna's creations.

In Figure III.57-5d, the watercolor illusion breaks the good continuation principle, eliciting new kinds of irregular shapes. Figure III.57-5e shows the watercolor illusion respectively pitted in favor or against the parallelism principle, whose role appears to be weak.

In Figure III.57-5f, the figure–ground effect of the watercolor illusion wins against the similarity principle (Pinna, 2005): the crosses emerge as a figure far more easily than the stars. As illustrated in Figure III.57-5g, the watercolor illusion induces a clear figural effect also in this extreme conditions where the watercolored crosses bulge as figures even on a background not uniform but articulated and differentiated in many colors patches. This is the opposite of what is expected. Finally, Figures III.57-1 to III.57-3, as previously described, demonstrate that the watercolor illusion is also stronger than the past experience principle.



Video III.57-2. Video by Dr. Bainglio Pinna's creations.

RELATIONSHIP BETWEEN THE COLORATION AND THE FIGURE–GROUND EFFECTS

The understanding of the relationship between coloration and figure–ground effects is one of the basic problems of vision science. It starts from the following fundamental question: Are color and figure–ground processes independent?

Coloration Without Figure–Ground Effect

Pinna and Reeves (2006) demonstrated that coloration without the figure-ground effect can be induced through a dichoptic composite stimulus, with the orange contour to one eye and the purple one to the other eye. Under this condition, the coloration is clearly perceived within the edges, but the figure-ground effect is totally absent. Under these conditions, the coloration is clearly perceived within the edges, but the figure–ground effect is totally absent.

Coloration without any figure–ground effect can be obtained with equiluminant juxtaposed contours, as demonstrated in Figure III.57-2h, which induces a clear coloration but the figure-ground effect is very weak or even absent: the figure–ground organization is flat and reversible. This is also the case of the central region of Figure III.57-4d, where neither the stars nor the crosses prevail as a figure against the other. Another case of coloration without figure–ground effect is illustrated in Figure III.57-3d, where a dense illusory smog fills the interspaces among circles and creates a cyan inset and a reddish surrounding coloration without figure–ground effect. In Figure III.57-6a, the three couples of purple-orange juxtaposed contours annuls the figure–ground and volumetric 3D properties of the watercolor illusion but not the coloration effect.

Figure–Ground Without Coloration Effect

We have already shown an example of figure–ground effect without coloration in Figure III.57-5f; the crosses clearly emerge as rounded convex figures while the inner coloration is about absent. The figure–ground without coloration effect is even stronger in Figure III.57-6b, where the inner edge of each cross appears as a dirty achromatic color, denser and brighter than the color within each star.

Figure III.57-6b clearly shows the surface capture due to the figural effect of the watercolor illusion. The gray within the crosses is perceived as captured in a more elevated depth plane than the gray within the word stars that is perceived instead as pushed into the background.

Another clear condition of the figure–ground effect without coloration is obtained by increasing the number of the juxtaposed contours, as illustrated in Figure III.57-6c.

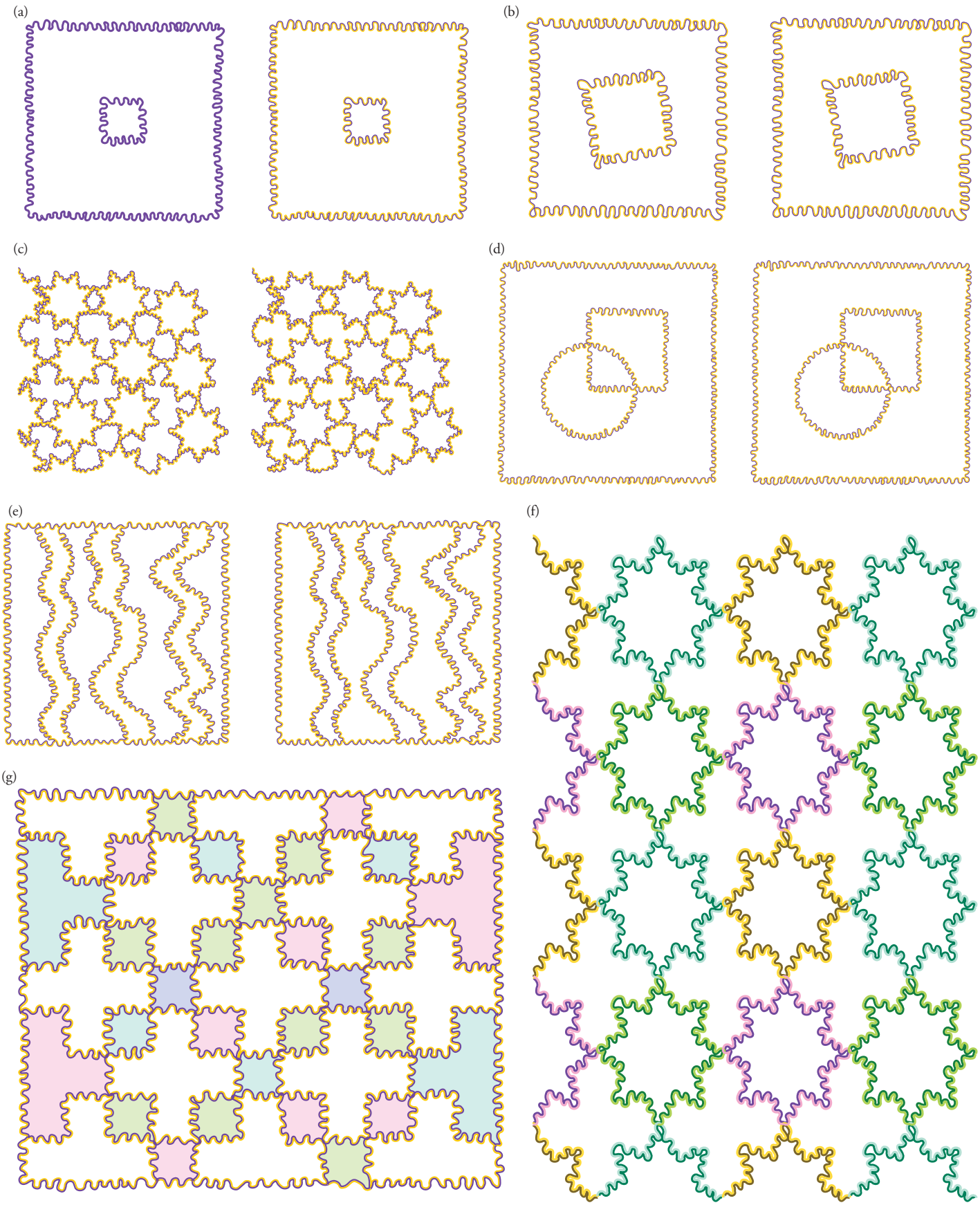


Figure III.57-5. The figure-ground effect of the Watercolor illusion. (a–g) The watercolor illusion wins against grouping and figure-ground principles: (a) surroundedness; (b) relative orientation and surroundedness; (c) regularity and Prägnanz; (d) good continuation; (e) parallelism; (f) similarity; (g) background articulation. (Figures by Dr. Baingio Pinna's creations.)

A brown, undulating shape with a 3D appearance illuminated from the front by a bright light is perceived. The coloration effect is absent and the inner region is seen as a bright white, purer than the white of the background, or as a bright light reflected by the 3D brown shape. This is called the lighting illusion (Pinna, 2006; Pinna & Reeves, 2006).

The figure–ground effect without coloration occurring in the lighting illusion is supported by another illusion related to the former. This is the discoloration illusion (Pinna, 2006) illustrated in Figures III.57-6d and III.57-6e, where the inner region of Figure III.57-6d appears white even if it is physically filled with the same gray of Figure III.57-6e. The gray is phenomenally discolored.

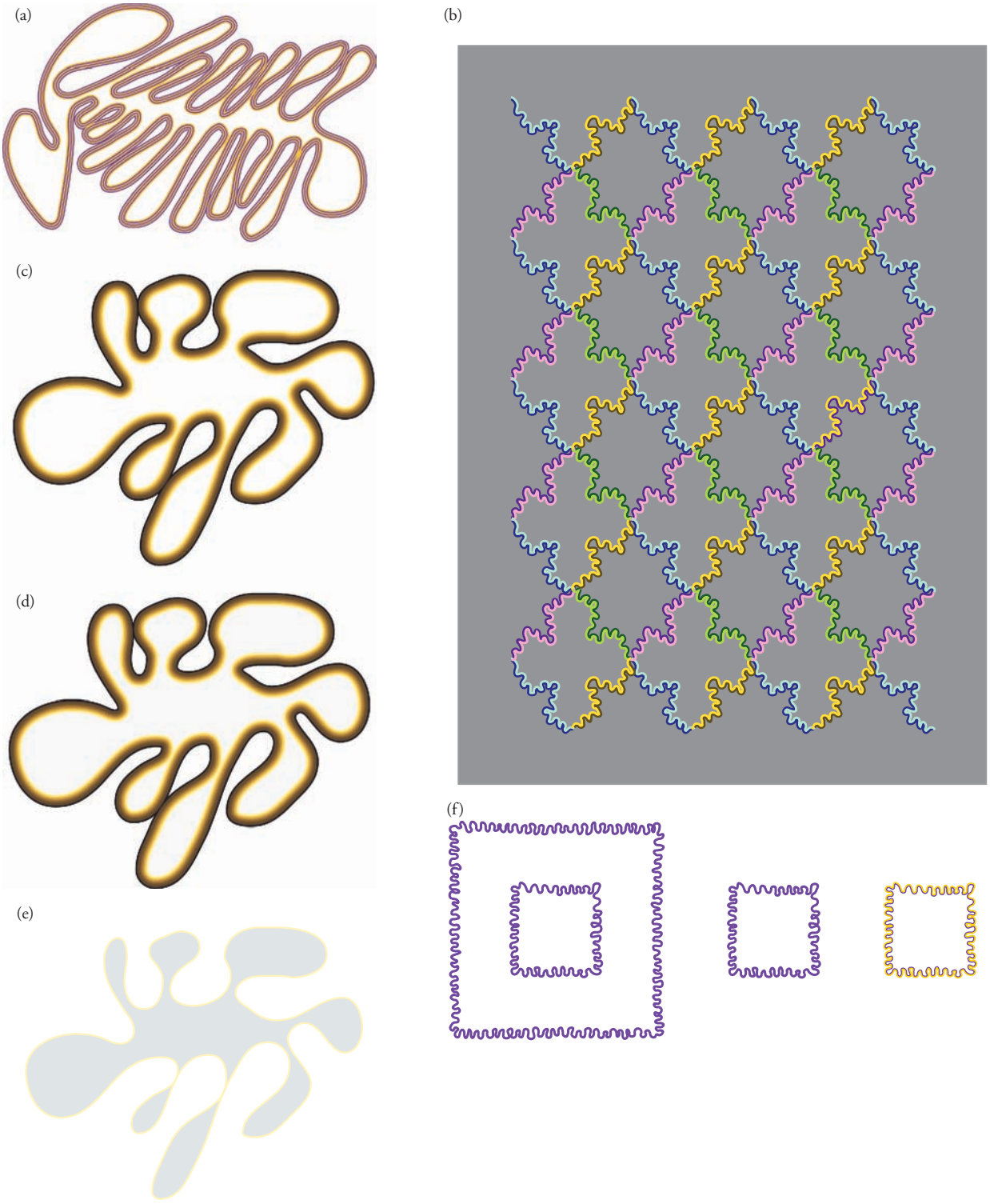


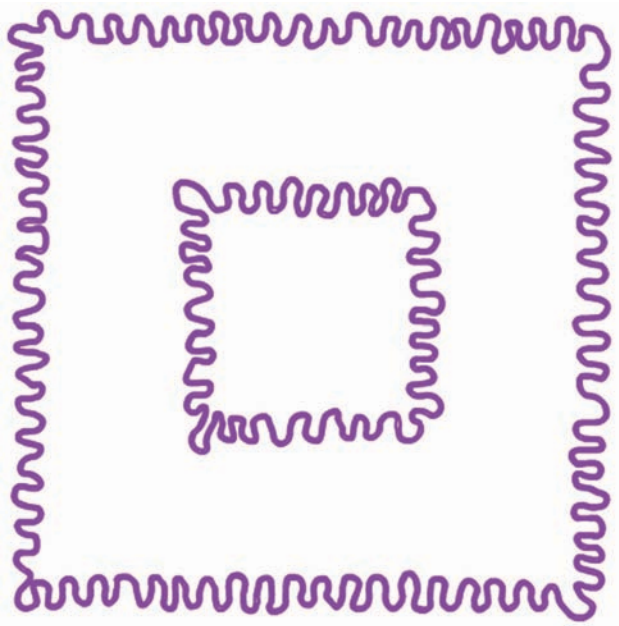
Figure III.57-6. (a) The figure-ground effect is annulled, while the coloration effect persists. (b) Surface capture due to the figural effect of the watercolor illusion. (c) The lighting illusion. (d-e) The discoloration illusion. (f) Through the watercolor illusion a hole alone, without a surrounding object, can assume a figure status. (Figures by Dr. Baingio Pinna's creations.)

Figure–Ground and Object-Hole Effects

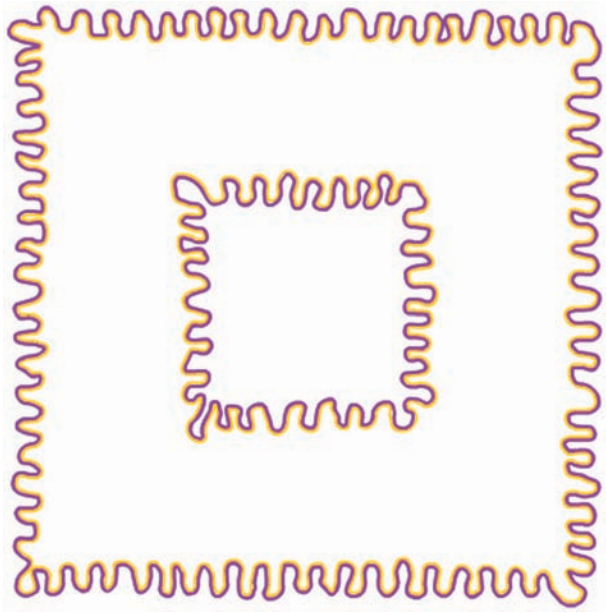
From the previous figures, it follows that the figure–ground effect of the watercolor illusion is more similar to an object-hole effect, that is, as the emergence of a volumetric object on the empty space perceived as a solid 3D hole (cf. Figures III.57-2a, III.57-5a, and III.57-5b).

The phenomenal notion of the hole induced by the watercolor illusion is more than something in between figure and background (see Bertamini, 2006; Palmer, 1999). It can be useful to solve the more general problem of object perception and shape formation. In Figure III.57-6f (left), the perception of the hole depends on the presence of the outer square surrounding it and appearing as the surface that contains the hole. The boundaries belong to the hole and at the same time to the complementary region. By removing the outer square, the hole becomes a figure (Fig. III.57-6f, center), thus losing the phenomenal properties of being a hole. This is not the case under watercolor conditions (Fig. III.57-6f, right), where even a hole alone (i.e., without a surrounding object) can assume a figure status. In Figure III.57-6f (right), the main result is a hole with the shape of an irregular square.

These results suggest several interesting conditions. By imparting motion to the inner square of Figure III.57-6f (left), the figure–ground segregation, eliciting the static hole, is expected to be broken. In fact, in Video III.57-3, by forcing the visual system to perceive the inner square shape as a hole under static conditions, when the inner square starts moving, the imparted motion induces a clear figural property. It appears as a small square moving within a large square. The question is now: If the watercolor illusion induces a strong figural property to the hole, can it still be perceived when motion is imparted? Video III.57-4 clearly demonstrates that the inner component is paradoxically perceived even more strongly as a hole. This result also occurs with a hole alone, as shown in Video III.57-5 and with the square-shape hole freely moving around the boundaries of the larger irregular square (Video III.57-6).



Video III.57-3. Video by Dr. Bainglio Pinna's creations.



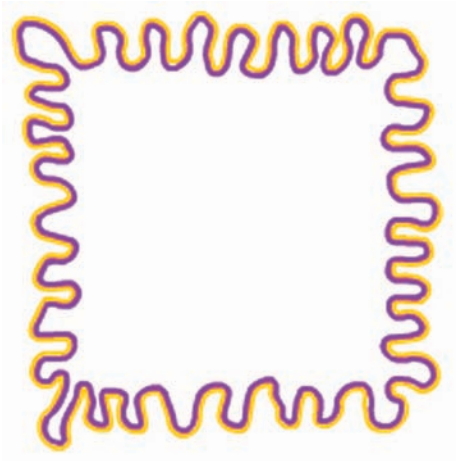
Video III.57-4. Video by Dr. Bainglio Pinna's creations.

The Asymmetric Luminance Contrast Principle

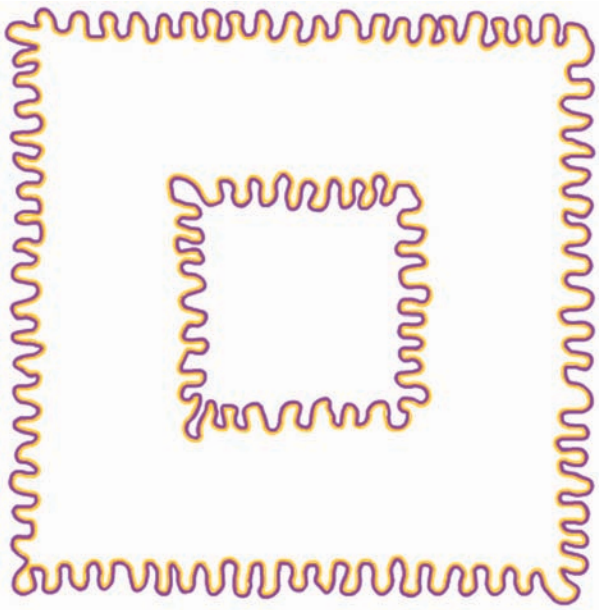
The figure–ground and object-hole effects discussed here suggest that the watercolor illusion can be considered the result of a new principle of grouping and figure–ground segregation that is much stronger than the Gestalt ones. This principle, called the asymmetric luminance contrast principle (Pinna, 2005), states that, all else being equal, given an asymmetric luminance contrast on both sides of a boundary, the region, whose luminance gradient is less abrupt, is perceived as a figure (object) relative to the complementary more abrupt region perceived as a background (hole).

NEURAL MECHANISMS UNDERLYING THE WATERCOLOR ILLUSION

The relative independence between the coloration and figural effects (Pinna & Tanca, 2008) suggests the existence of



Video III.57-5. Video by Dr. Bainglio Pinna's creations.



Video III.57-6. Video by Dr. Bainglio Pinna's creations.

two parallel but related mechanisms. The FACADE model (Grossberg, 1994, 1997, 2000; Grossberg & Todorović, 1988) posits that two processes, boundary grouping and surface filling-in (Komatsu, 2006; Murakami, 2008) substantiated by the cortical interblob and blob streams, respectively, within cortical areas V1 through V4, can explain the coloration and the figure–ground effects in the watercolor illusion (Grossberg & Mingolla, 1985a, 1985b; Grossberg & Todorović, 1988; Pinna & Grossberg, 2005). These boundary and surface processes show complementary properties, and, through their interaction, they generate a consistent perceptual representation that overcomes the complementary deficiencies of each stream, acting on its own. Boundary and surface processes are modeled by the boundary contour system and the feature contour system (Grossberg & Mingolla, 1985a, 1985b).

Other findings (Friedman, Zhou, & von der Heydt, 2003; von der Heydt, Zhou, & Friedman, 2003; Zhou, Friedman, & von der Heydt, 2000) show that neurons in V2 respond with different strength to the same contrast border, depending on the side of the figure to which the border belongs, implying a neural correlate process related to the unilateral belongingness of the boundaries. According to these results, figure–ground segregation may be processed in areas V1 and V2, in inferotemporal cortex and the human lateral occipital complex. Not only the figure–ground segregation but also the color tint of the watercolor illusion might have its explanation in the cortical representation of borders (von der Heydt & Pierson, 2006).

The differentiated and specialized figural roles of the juxtaposed contours and of their colors suggest a possible neural scenario where multiple juxtaposed lines may stimulate neurons, selective for different asymmetric luminance profiles, signaling not only the unilateral belongingness of the boundaries but also the coloration. Due to the gradient variations at different scales, it can be suggested that these neurons, at the beginning undifferentiated, become more and more specialized by assuming different roles but, at the same time, becoming part of the more global integrated process that codes the figurality due to the binding between figure–ground and coloration.

ACKNOWLEDGMENTS

The author was supported by Banco Fondazione di Sardegna - Rep. n. 2539. Prot. n. 23907 del 5/10/2015.

REFERENCES

Bertamini, M. (2006). Who owns the contour of a visual hole? *Perception*, *35*, 883–894.

Broerse, J., & O'Shea, R. P. (1995). Local and global factors in spatially contingent coloured aftereffects. *Vision Research*, *35*, 207–226.

Broerse, J., Vladusich, T., & O'Shea, R. P. (1999). Colour at edges and colour spreading in McCollough effects. *Vision Research*, *39*, 1305–1320.

Devinck, F., Delahunt, P. B., Hardy, J. L., Spillmann, L., & Werner, J. S. (2005). The watercolor effect: Quantitative evidence for luminance-dependent mechanisms of long-range color assimilation. *Vision Research*, *45*, 1413–1424.

Devinck, F., Delahunt, P. B., Hardy, J. L., Spillmann, L., & Werner, J. S. (2006). Spatial dependence of color assimilation by the watercolor effect. *Perception*, *35*, 461–468.

Devinck, F., Hardy, J. L., Delahunt, P. B., Spillmann, L., & Werner, J. S. (2006). Illusory spreading of water-color. *Journal of Vision*, *6*(5), 7, 625–633.

Devinck, F., & Knoblauch, K. (2012). A common signal detection model accounts for both perception and discrimination of the watercolor effect. *Journal of Vision 12*(3), 19, 1–14.

Devinck, F., & Spillmann, L. (2009). The watercolor effect: Spacing constraints. *Vision Research*, *49*, 2911–2917.

Devinck, F., Spillmann, L., & Werner, J. S. (2006). Spatial profile of contours inducing long-range color assimilation. *Visual Neuroscience*, *23*, 573–577.

Friedman, H. S., Zhou, H., & von der Heydt, R. (2003). The coding of uniform color figures in monkey visual cortex. *Journal of Physiology*, *54*, 593–613.

Grossberg, S. (1994). 3-D vision and figure–ground separation by visual cortex. *Perception & Psychophysics*, *55*, 48–120.

Grossberg, S. (1997). Cortical dynamics of three-dimensional figure–ground perception of two-dimensional pictures. *Psychological Review*, *104*, 618–658.

Grossberg, S. (2000). The complementary brain: Unifying brain dynamics and modularity. *Trends in Cognitive Sciences*, *4*, 233–245.

Grossberg, S., & Mingolla, E. (1985a). Neural dynamics of form perception: Boundary completion, illusory figures and neon color spreading. *Psychological Review*, *92*, 173–211.

Grossberg, S., & Mingolla, E. (1985b). Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations. *Perception & Psychophysics*, *38*, 141–171.

Grossberg, S., & Todorović, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena. *Perception & Psychophysics*, *43*, 241–277.

Komatsu, H. (2006). The neural mechanisms of perceptual filling-in. *Nature Reviews/Neuroscience*, *7*, 220–231.

Murakami, I. (2008). Perceptual filling-in. In M. D. Binder, N. Hirokawa, & U. Windhorst (Eds.), *Encyclopedia of neuroscience*. Berlin: Springer.

Palmer, S. E. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA: MIT Press.

Pinna, B. (1988). Un effetto di colorazione. In V. Majer, M. Maeran, & M. Santinello (Eds.), *Il laboratorio e la Città. XXI Congresso*

degli Psicologi Italiani (p. 158). Palermo, Italy: Società italiana di psicologi.

Pinna, B., (2005). The role of the Gestalt principle of similarity in the watercolor illusion. *Spatial Vision*, *21*, 1–8.

Pinna, B. (2006). The discoloration illusion. *Visual Neuroscience*, *23*, 583–590.

Pinna, B. (2008a). A new perceptual problem: The amodal completion of color. *Visual Neuroscience*, *25*, 415–422.

Pinna, B. (2008b). The watercolor illusion. *Scholarpedia*, *3*(1), 5352.

Pinna, B., Brelstaff, G., & Spillmann, L. (2001). Surface color from boundaries: A new "watercolor" illusion. *Vision Research*, *41*, 2669–2676.

Pinna, B., & Grossberg, S. (2005). The watercolor illusion and neon color spreading: A unified analysis of new cases and neural mechanisms. *Journal of the Optical Society of America A*, *22*, 2207–2221.

Pinna, B., & Reeves, A. (2006). Lighting, backlighting and watercolor illusions and the laws of figurality. *Spatial Vision*, *19*, 341–373.

Pinna, B., & Tanca, M. (2008). Perceptual organization reconsidered in the light of the watercolor illusion: The problem of perception of holes and the object-hole effect. *Journal of Vision*, *8*(7), 8, 1–15.

Pinna, B., Werner J. S., & Spillmann, L. (2003). The watercolor effect: A new principle of grouping and figure–ground organization. *Vision Research*, *43*, 43–52.

Rubin, E. (1921). *Visuell wahrgenommene Figuren*. Kobenhavn: Gyldendalske Boghandel.

Tanca, M., & Pinna, B., (2008). The phenomenal dissociation between coloration and object-hole effects in the watercolor illusion. *Visual Neuroscience*, *25*, 423–432.

von der Heydt, R., & Pierson, R. (2006). Dissociation of color and figure–ground effects in the watercolor illusion. *Spatial Vision*, *19*, 323–340.

von der Heydt, R., Zhou, H., & Friedman, H. S. (2003). Neural coding of border ownership: Implications for the theory of figure–ground perception. In M. Behrmann, R. Kimchi, & C. R. Olson (Eds.), *Perceptual organization in vision: Behavioral and neural perspectives* (pp. 281–304). Mahwah, NJ: Lawrence Erlbaum.

Werner, J. S., Pinna, B., & Spillmann, L. (2007). The brain and the world of illusory colors. *Scientific American*, *3*, 90–95.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt II. *Psychologische Forschung*, *4*, 301–350.

Zhou, H., Friedman, H. S., & von der Heydt, R. (2000). Coding of border ownership in monkey visual cortex. *Journal of Neuroscience*, *20*, 6594–6611.

Chapter 58
The Chinese Lantern Illusion

Sergio Roncato, Sandro Bettella, and Clara Casco

The apparent distortion of a straight line is one of the more frequently observed optical illusions, and the Zöllner and Orbison illusions are well-known examples of this visual effect. Any straight line can, nevertheless, seem to be distorted or wavy when there are local tilt effects, such those generated in the café wall or Fraser illusions.

In this chapter we are concerned with a distortion in the perception of straightness not with reference to a line but to a contour formed by a sequence of aligned angles, such as those forming the borders of Chinese lanterns, as depicted in Figure III.58-1.

The four objects depicted in this figure are equally sized, vertically aligned diamonds resembling Chinese lanterns. The illusion under concern here has thus been called the "Chinese lantern illusion." Figures III.58-1b, III.58-1c, and III.58-1d are similar, but only Figure III.58-1d appears as it really is. The vertical margins of the other two objects seem to expand (Fig. III.58-1b) or to contract (Fig. III.58-1c). The illusion is even stronger in Figure III.58-1a as the vertical borders seem to form a waving outline. It is instead identical to the lantern in Figure III.58-1c, the only difference being that while the jagged borders of Figure III.58-1c are of a uniform grey color, they are varishaded grey with luminance increasing and decreasing.

Figure III.58-1 provides important information that helps to explain how the illusion is created. There is no

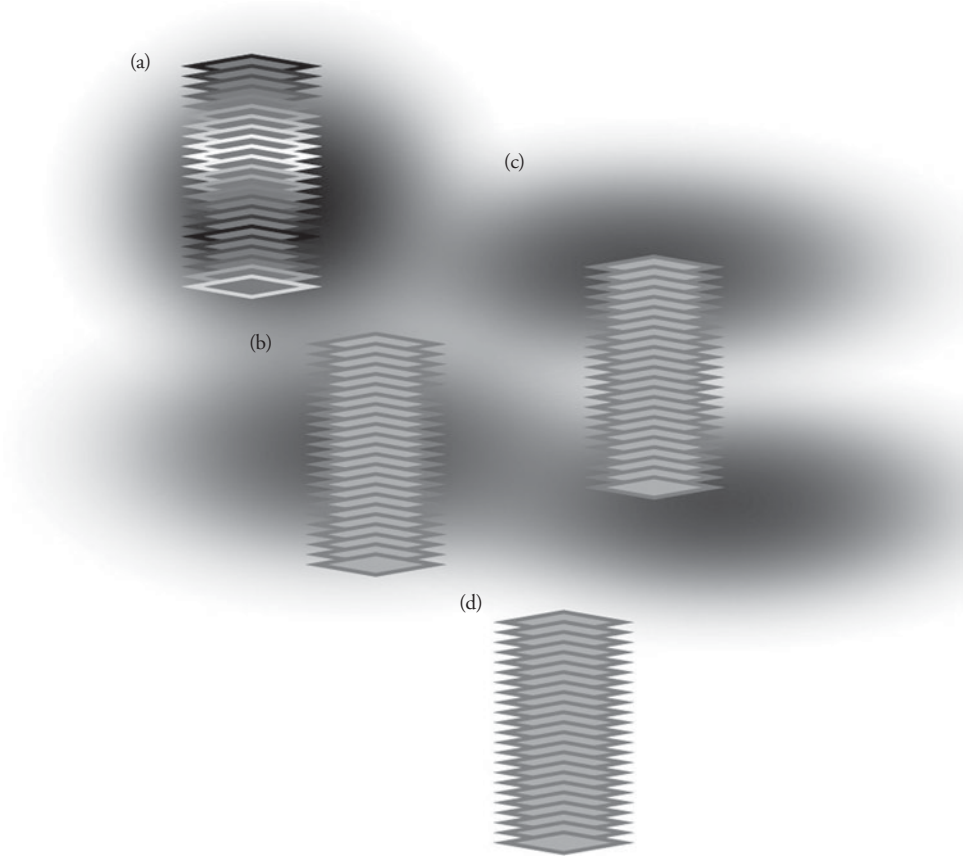


Figure III.58-1. Four Chinese lanterns are made up of equal-sized diamonds aligned in columns. The impression that three of them are bulging or shrinking is illusory. (Authors own unpublished work.)

distortion of the vertical borders when the background luminance is homogeneous. A distortion of the borders becomes evident in those cases in which the background luminance varies, becoming lighter or darker.

The phenomenon seems to belong to a class of illusions in which the location and straightness of the contour is altered when one or more of the surfaces (the figure or the background) or the outline borders undergo luminance modulation in space (luminance gradient or luminance ramp) or in time. Anstis and Rogers (1975) and Gregory and Heard (1983) demonstrated that a rectangle outlined by a border that is lighter on one side and darker on the other appears to be displaced from its real location in a direction depending upon the background luminance. Shapiro, Charles, and Shear-Heyman (2005) took these findings a step further, demonstrating that even the luminance modulation of a contour in space and in time can affect how figures are perceived. They demonstrated illusory effects in complex figures in which local distortion effects work together to create an overall tilting or movement effect.

Observing illusory distortions of an outline contour separating two surfaces, one of which having luminance varying in space or in time, Van Lier and Csathó (2006) arrived at the conclusion that the divergence of results in temporal and spatial domains is apparent because the contour deviation can be predicted by the same rule: the apparent border displacement tilts in the direction that preserves contrast polarity across edges separating a border from a uniform grey surface. Kitaoka (2004) also described an impressive illusory motion effect in static stimuli made up of ellipsoids contoured on one side by a black outline and on the other side by a white border.

In a previous work (Roncato, 2000; Roncato & Casco, 2003), we analyzed illusory deviations from straightness of thin contours when portions of the surfaces on opposite sides vary in luminance or when the border is colored by varying shades of grey. Two of these situations are illustrated in Figure III.58-2c–d. A luminance ramp is interposed between two vertical columns of different shades of grey. The vertical border is interrupted periodically where the luminance of the adjacent background is similar or the same and then continues with a reversed-contrast polarity (see figure caption). We hypothesized that the perceptual prolongation takes a different direction: it deviates to conjoin with the nearest border having the same contrast polarity (the CP rule, see Fig. III.58-2 caption). This local tilt is the basic effect, causing a tilting or waving effect. Van Lier and Csathó (2006) outlined some illusory effects not predicted by this rule and requiring a theoretical framework in which global as well local brightness differences play a role.

It is possible, then, on the basis of these considerations, to classify the bulging/shrinking lantern illusion as a manifestation of well-known displacement effects due to a luminance variation. But the fact that in this case the contour being examined is formed by a series of triangular waves requires further analysis.

There are two pairs of bands in Figure III.58-2a–b. Each part of the figure is made up of a vertical band with sawtooth borders and to its right a second vertical band formed

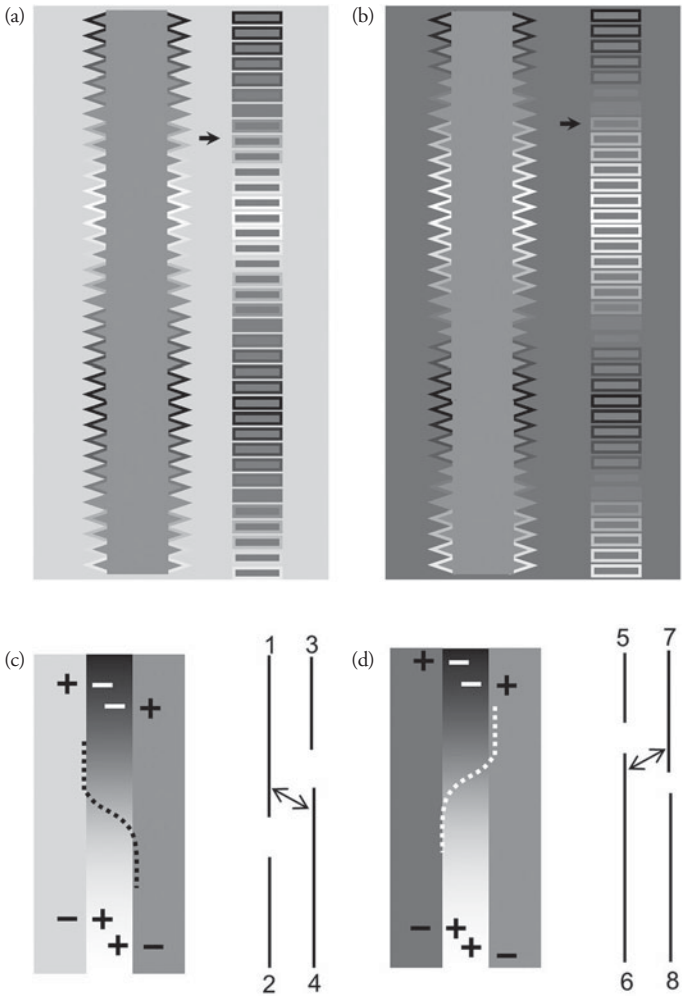


Figure III.58-2. (a) Of the two grey vertical bands, the band on the left is bordered on both sides by triangles contoured by an outline that is cyclically modulated in shades of grey (from black to white and vice versa). The band on the right is made up of a series of aligned, touching, uniformly colored grey rectangles outlined by lighter or darker shades of grey matched to the triangles in the band on the left. (b) Differs only in that the grey background against which the two bands are depicted is much darker with respect to that in b. (c) and (d) A magnified surface in the portion of space in which the outline contour varies in luminance from black to white. The rectangles (columns) at the opposite sides of the luminance gradient reproduce the luminances of the backgrounds (left) and of the figures (right). "+" and "–" indicate the lighter and the darker surfaces, respectively, at the opposite sides of an edge. Numbered vertical lines symbolized the edges. The gap separating them (i.e., the region of equiluminance) can be filled by a prolongation along a straight direction (completion of the collinear couples 1–2, 3–4, 5–6, and 7–8) that requires the contrast polarity reversal. The CP rule predicts an alternative completion if edges of the same contrast polarity are nearby. It is symbolized by a bidirectional arrow indicating the couples of edges 1–4 and 6–7 as candidates to completion in the (c) and (d) conditions, respectively. The completion following this direction is assumed to prevail on concurrent ones that, though preserving the contrast polarity, fill longer gaps: conjoining the edges 2–3 and 5–8 requires a longer link than acting in the opposite directions 1–4 and 6–7. The dotted curved line traversing the luminance gap indicates the predicted "path of conjunction" of the edges with the same contrast polarity. (Authors own unpublished work.)

by small vertically aligned rectangles. The sawtooth border is formed by equally sized triangles, and the external points of the triangle form an illusory border with cyclic luminance modulations. Even the borders of the aligned rectangles in the figure are variable. The variations follow the same pattern in both parts of the figure in that the areas where the sawtooth border are darker correspond horizontally to the areas where the rectangles are darker. In other words, the luminance profile of the two outline contours corresponds to the same waveform.

It is important to point out that even the illusory oscillation is synchronized. In Figure III.58-2a, the two bands seem to expand where the border is darker and to contract where it is lighter. The same synchronicity can be seen in Figure III.58-2b, but here the band seems to expand in the area where the border is lighter.

With regard to the aligned rectangles, the directions of contour distortion are predictable on the basis of the CP rule, as Figures III.58-2c and III.58-d illustrate. The central luminance ramp and the lateral bands reproduce the luminance conditions at the borders of the aligned rectangles. The dashed curved line symbolizes the prolongation of the borders preserving the contrast polarity and makes evident the local tilting that generates a waving impression (see Fig. III.58-2 caption). Small arrows in Figure III.58-2a–b indicate the region where the deviation predicted in Figure III.58-2c–d occurs. Thus the paths along which the edges seem to be prolonged predict the direction of the perceived orientation tilt.

We can make three interesting observations about Figure III.58-2. First, since, as has been pointed out, the illusory oscillations in the bands are synchronous, it can be hypothesized that this effect is caused by the same basic mechanism. The sawtooth border seems to behave as if it were undergoing the same distorting effect described in Figures III.58-2c and III.58-2d.

Second, the shapes forming the parts of the figures—that is, the triangles forming the borders of the band on the left and the rectangles on the right—seem to be the same size and are separated by uniform spaces. These features are not congruent with alternative interpretations of the effect in terms of border displacement (Gregory & Heard, 1983: Shapiro et al., 2005; Shapiro & Knight, 2008). According to those theories, when a contour bulges, the shapes being outlined seem to expand, so the triangles and rectangles in this figure should appear to be larger when, in fact, they do not.

Third, the amplitude of the oscillation of the sawtooth contours in Figure III.58-2 is larger than that observed for the pile of aligned rectangles; this leads to the hypothesis that in the former case if the CP rule is applicable, its influence seems to go beyond the limits of the short spatial range postulated by previous research (Roncato & Casco, 2006).

How can these contrasting data be reconciled? The jagged contours in Figures III.58-2a and III.58-2b seem to behave like thick contours, and the CP rule appears to be acting within a larger spatial range.

Let us turn our attention to a configuration that may provide more clues to the solution. In Figure III.58-3, the luminance gradient has a rectangular shape. Figure III.58-3a is made up of a variable grey-colored band (with the lightest shades at the far ends) lined at its two horizontal borders with rows of small, upright, uniformly colored grey rectangles. The band is depicted against a uniform, light-grey background. Figure III.58-3b is identical to Figure III.58-3a except for the fact that the background is a darker shade of grey.

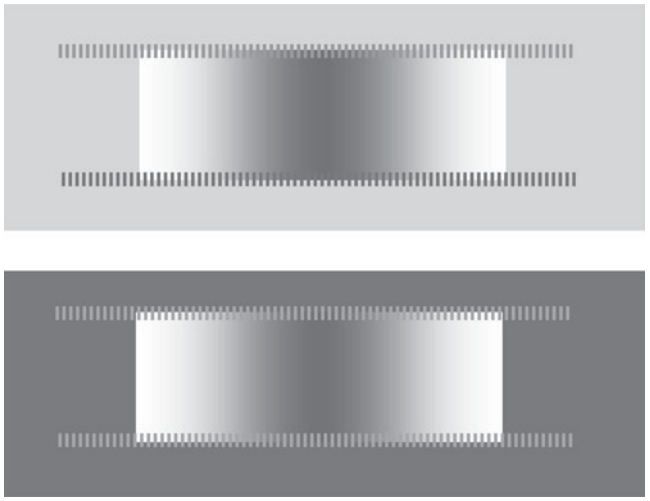


Figure III.58-3. (a) A variable grey-colored band (with the lightest shades at the far ends) lined at its two horizontal borders with rows of small upright, uniformly colored grey rectangles. The band is depicted against a uniform, light-grey background. (b) Identical (a) except for the fact that the background is a darker shade of grey. This variation has important consequences on the direction of horizontal borders distortion that appear to bulge outward. (Authors own unpublished work.)

The impression is that the central part of the horizontal band in Figure III.58-3a is contracting. If we observe the band's horizontal borders, the reason for this effect becomes apparent. We can see that they separate the dark internal surface area from the lighter background. If we trace one of these borders, for instance the top one, going right at a certain point, it seems to vanish behind the rows of vertical rectangles. We have the impression that it is continuing to unite itself with the illusory border formed by the upper extremities of the rectangles. If we trace the border from the center and go left, the effect is the same—apparent but symmetric—explaining why the top border seems to contract. The contraction toward the lower border is produced by symmetric effects on those generated by the upper border.

Thus the concavity effects seem to result from a "hybrid" conjunction of a continuous border and the illusory border formed by the outer extremities of the thin small rectangles. The impression of convexity in Figure III.58-3b is the consequence of the hybrid conjunction of the extreme black/grey edges with the central extremities of the small rectangles.

What are the properties of this "path of conjunction"? It is clearly visible in those areas in which there is a high level of luminance contrast and appears to extend into the lateral extremities of segments that preserve the contrast polarity.

If the assumption is made that the acute angles in Figure III.58-1 are segments attached to a luminance

ramp, this should lead to the same illusory phenomena. In Figure III.58-4a there is a variable grey-colored band, darkest on the left lateral side and lightest on the right. The band is horizontally bordered above and below by rows of small, vertical rectangles all the same shade of grey and beginning before and ending after the central band. The band is set against a uniform medium dark-grey background. In Figure III.58-4b there is a series of aligned equal-sized diamonds in darker to lighter shades (going from left to right) of grey outlined by borders all in the same grey color. The figure (a Chinese lantern lying on its side) is depicted against a medium-grey background (the same as the one in Fig. III.58-4a). In Figure III.58-5a there is a uniform grey horizontal band lined at the upper and lower horizontal borders by rows of small, upright rectangles periodically becoming a darker or lighter grey. In Figure III.58-5b a grey horizontal band is contoured at its horizontal borders by a sawtooth outline periodically varying in luminance. These figures allow us to compare the illusory effects generated by contouring rows of segments or jagged outlines.

In Figures III.58-4a and III.58-4b, the illusory expansion/contraction is synchronous; that is, the central horizontal area appears to bulge at the dark side. A synchronous oscillation is perceived also in Figures III.58-5a and III.58-5b; in fact the illusory bulging is perceived wherever the borders (rectangles or triangles) are darkest. In the latter case the bulging and the shrinking alternate periodically, seeming to oscillate in an imaginary waving line highest at the peaks of the darkest triangles and lowest at the valleys of the white ones. This line (a portion of which corresponds to the dashed line) has two properties: it conjoins the borders with the highest contrast magnitudes, and it preserves the contrast polarity throughout.

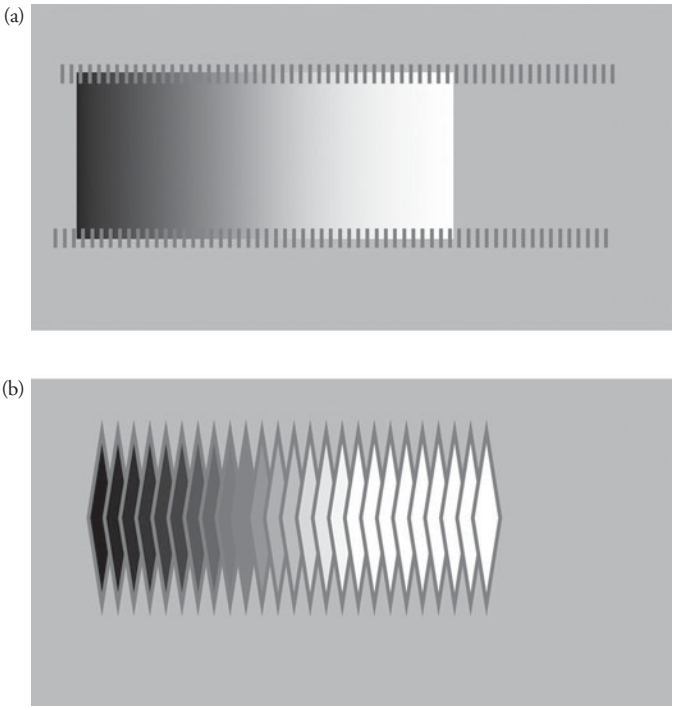


Figure III.58-4. (a) Variable grey-colored band, darkest on the left lateral side and lightest on the right lateral side. The band is bordered above and below horizontally by rows of small, upright grey rectangles beginning before and ending after the central band. The figure is set against a uniform medium-grey background. (b) A series of aligned equal-sized diamonds going from the darkest grey on the right to the lightest grey (white) on the left. The borders of the diamonds are all in the same grey shade. The figure (a Chinese lantern lying on its side) is set against the same medium-grey background as that in (a). (Authors own unpublished work.)

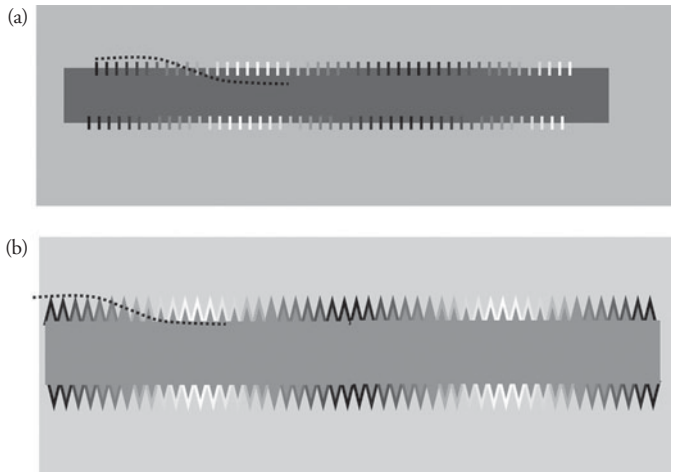


Figure III.58-5. (a) A uniform grey horizontal band lined at the upper and lower horizontal borders by rows of small, vertical variably colored light- to dark-grey rectangles beginning before and ending earlier with respect to the central band. The dark, curved dotted line at the far left end of the figure indicates the direction of the perceived oscillation of the border. Note that this line is tangent to the extremities with the greatest amount of luminance in contrast to the background. (b) A central uniform grey band lined at its upper and lower horizontal borders by grey triangles whose outline contour forms a sawtooth line cyclically varying in luminance. The curved, dark, dotted line at the far left of the band indicates the direction of the perceived oscillation of the border. Note that this line is tangent to the extremities with the greatest amount of luminance contrast to the background. (Authors own unpublished work.)

The same characteristics can be found in Figure III.58-4a depicting an illusory waving contour. The starting point of the path is the highly contrasted black/light grey contour on the left and the ending point is the illusory border—with the same contrast polarity—formed by the outer extremities of the line of small, upright rectangles.

The hypothesis that the vertices of acute angles can behave as lines and generate a waving impression conflicts with findings that distortions are found only in straight lines. The impression of oscillation is, instead, perceptually vivid even in this case with reference to jagged contours.

The phenomenon is similar to that observed by the Shapiro and Knight (2008, Figure 8c) with reference to a pattern of blurred disks against a luminance gradient: thick, blurred contours appeared to deviate from straightness, "drawing" a path in which the contrast polarity was preserved. This suggests that the perception of fragments merging is more general; that is, the visual system is likely to join a class of borders, even if they are blurred and/or jagged, following the contrast polarity rule (Parlangeli & Roncato, 2008; Roncato & Casco, 2009).

AUTHORS' NOTE

Presented at the Best Illusion of the Year contest, the "Skyscaper and Clouds" illusion was one of the 10 to be awarded in 2008 (http://illusioncontest.neuralcorrelate.com/cat/top-10-finalists/2008/#post-132). It can be seen in the animation showing skyscrapers and clouds depicted against a cloudy sky. The illusory bulging/contraction effect is evident when the jagged dark- or light-grey contours of the buildings and skyscrapers are projected against a variable sky, at times clear and bright and at times animated by variably colored grey clouds shooting upward.

The lantern illusion described in this chapter is an illustration of a luminance ramp effect.

REFERENCES

Anstis, S. M., & Rogers, B. J. (1975). Illusory reversals of movement and depth during changes in contrast. *Vision Research*, *15*(8), 957–961.

Gregory, R., & Heard P., (1983). Visual dissociations of movement, position, and stereo depth: Some phenomenal phenomena. *Quarterly Journal of Experimental Psychology A, 35*(1), 217–237.

Kitaoka, A. (2004). The latest works 2. http://www.psy.ritsumei.ac.jp/~akitaoka/saishin2e.html

Lier, R. van, & Csathó, Á. (2006). Dancing shapes: A comparison of luminance-induced distortions. *Perception*, *35*(6), 775–798.

Parlangeli, O., & Roncato, S. (2008). Boundary completion, contrast polarity, and the perception of illusory tilt. *Perception, 37*(4), 535–556.

Parlangeli, O., & Roncato, S. (2010). Draughtsmen at work. *Perception, 39*(2), 255–259.

Roncato, S. (2000). The effect of luminance variation on the apparent position of an edge. *Attention, Perception & Psychophysics*, *62*(4), 762–785.

Roncato, S., & Casco, C. (2003). The influence of contrast and spatial factors in the perceived shape of boundaries. *Attention Perception & Psychophysics*, *65*(8), 1252–1272.

Roncato, S., & Casco, C. (2006). Illusory boundary interpolation from local association field. *Spatial Vision* , *19*(6), 581–603.

Roncato, S., & Casco, C. (2009). A new "tilt" illusion reveals the relation between border ownership and border binding. *Journal of Vision*, *9*(6), 14, 1–10. http://jov.arvojournals.org/article.aspx?articleid=2204008

Shapiro, A. G., Charles, J. P., & Shear-Heyman, M. (2005). Visual illusions based on single-field contrast asynchronies. *Journal of Vision*, *5*(10), 764–782. http://journalofvision.org/5/10/2/

Shapiro, A., & Knight, E. (2008). Spatial and temporal influences on the contrast gauge. *Vision Research, 48*(6), 2642–2648.

Chapter 59
The Wedding Cake Illusion

Interaction of Geometric and Photometric Factors in Induced Contrast and Assimilation

Branka Spehar and Colin W. G. Clifford

DESCRIPTION OF THE EFFECT

The perception of surface lightness is one of the most immediate and fundamental aspects of human vision, often trivialized for its apparent simplicity. However, the relationship between the physical characteristics of light reflected from a surface and its perceived qualities is complex and remains only partially understood. The appearance of a surface as white, grey, or black is determined not only by the radiant intensity of that surface but also by the relative radiance of surrounding regions. The study of these interactions, known as induction effects, has a long-standing and important status in the history of vision research, as there is no attribute of a region of visual space that is not in a simple or more complex way influenced by its spatial context. These ubiquitous contextual influences in lightness perception are broadly classified as either contrast or assimilation, depending on whether the induced changes in surface appearance are shifted *away from* or *toward* the surrounding context, respectively.

The majority of research on context-dependent interactions in surface appearance has traditionally focused on contrast effects. In one of the most famous examples, simultaneous lightness contrast (Fig. III.59-1a), a grey test region of equal physical luminance appears lighter when viewed on a dark surround and darker when viewed on a light surround (Heinemann, 1955). The most common explanation for contrast effects are inhibitory interactions between retinal or cortical neurons, functioning to enhance differences in luminance or color and to aid in object detection and surface segregation. These mechanisms, however, are generally limited in their ability to explain assimilation effects, where the perceived lightness or color of the surface is shifted not away from but toward the characteristics of the inducing context. One of the earliest demonstrations of assimilation is the von Bezold effect (Fig. III.59-1b) in which superimposed white elements lighten and superimposed black elements darken the contiguous gray background area (Albers, 1975).

Assimilation was initially attributed to the "spreading" and optical mixing of stray light, a notion supported by findings of Helson (1963), who found that the relative width of the superimposed stripes determined whether assimilation or contrast ensued. Assimilation was observed with lines up to 8 to 10 mm wide after which classical contrast occurred, revealing a continuum from assimilation to contrast. These early findings led to the conclusion that assimilation occurs when the inducing regions are small relative to the test region, while contrast results when the inducing regions are large relative to the test region. However, since then, assimilation effects have been found across a wide range of spatial configurations (see Fig. III.59-1c–d), and effects of this type remain puzzling and particularly challenging for theories of lightness perception (Agostini & Galmonte, 2002; Bindman & Chubb, 2004; DeValois & DeValois, 1988; De Weert & Spillman, 1995; Hong & Shevell, 2004; Shapley & Reid, 1985). It also continues to be debated whether "contrast" and "assimilation" reflect different underlying mechanisms or are just a convenient way to describe superficial differences in the direction of various contextual effects (Kingdom, 2011).

One of the most striking and relatively more recent examples of assimilation is White's effect (White, 1979), where physically identical grey patches placed on the black and white bars of a square-wave grating appear markedly different in lightness (Fig. III.59-2a). The grey patches on the white bars, more extensively bounded by black borders than by white ones, appear darker than the grey segments on the black bars. If the total induced effects were simply the sum of contrast effects induced by the light and dark parts of the surround, one would expect the grey segments on the white bars to appear lighter, but that is most definitely not the case. Numerous attempts have been made to explain White's effect, ranging from early-stage cortical-filtering operations (Blakeslee & McCourt, 1999; Foley & McCourt, 1985; Kingdom & Moulden, 1991; Moulden & Kingdom, 1989) to higher-level perceptual interpretations based on depth or perceived belongingness (Anderson, 1997, 2003; Gilchrist et al., 1999; Ross & Pessoa, 2000; Spehar, Clifford & Agostini, 2002; Todorović, 1997; Zaidi, Spehar, & Shy, 1997). Some of the most popular accounts of White's effect suggest that local depth cues in the image, such as T-junctions, act to promote differential grouping between the test patch and the surrounding regions, which in turn selectively inhibit or enhance contrast induced from these regions (Gilchrist et. al., 1999; Ross & Pessoa, 2000;

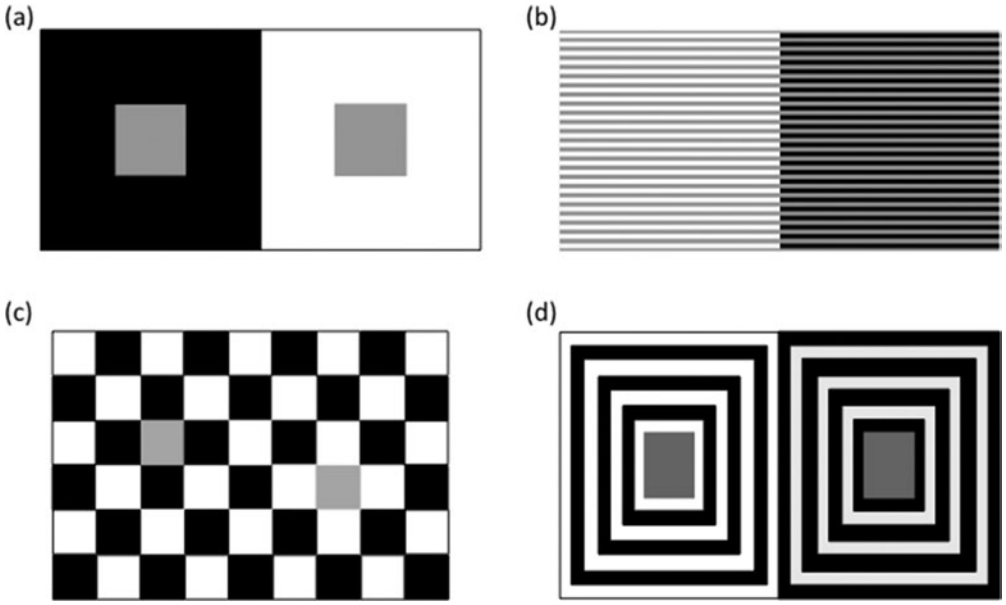


Figure III.59-1. Lightness contrast and lightness assimilation. (a) Simultaneous contrast effect: the two central grey squares are of equal physical luminance but the square surrounded by white appears darker than the square surrounded by black (Figure adapted from Hering [1964]). (b) Von Bezold effect (Figure adapted from Albers [1975]): the grey background area appears darker with superimposed black and lighter with superimposed white elements. (c) The checkerboard contrast illusion (Figure adapted from de Valois & de Valois [1988]): the two grey squares are physically identical, but the grey square surrounded predominantly with white checks appears lighter than the grey square surrounded predominantly by black checks. (d) The bullseye illusion (Figure adapted from Bindman & Chubb [2004]): the central grey square bordered by black is perceived as darker than the central square bordered by white.

Todorović, 1997; Zaidi et al., 1997). T-junctions in an image are local depth signs, typically interpreted as indicating occlusion by an opaque surface (Fig. III.59-2b). Given that the occluding surface typically corresponds to the region at the top of the T-junction, the grey patches are grouped more strongly with the inducing stripes they interrupt as opposed to the inducing stripes on the top and the bottom (which could be perceived as locally occluding each of the test patches). As a consequence, the contrast induced from these flanking top and bottom regions is inhibited compared to the contrast induced from the regions perceived to be coplanar with the test patches. Another T-junction–based model has been proposed by Anderson (1997), who argues that the effect is due to perceptual scission determined by the consistent luminance and contrast polarity relationships of aligned contours at T-junctions. In this model, the luminance of the target region is split into multiple sources based on the consistent contrast polarity relationships at all T-junctions for test regions interrupting black and white stripes.

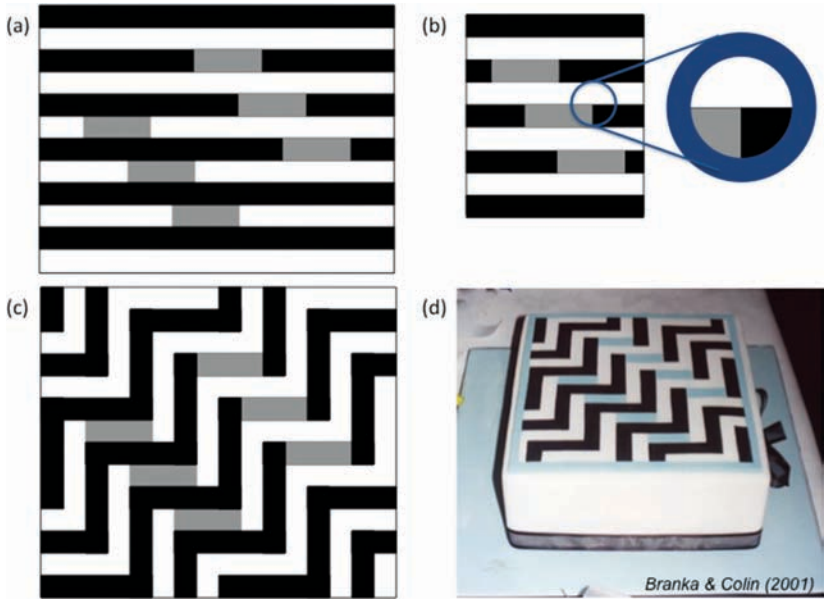


Figure III.59-2. (a) White's illusion (Figure adapted from White (1979)). (b) T-junction structure for one of the grey target areas in White's configuration: T-junctions associated with the grey target patches interrupting black stripes are consistent with an occlusion by a white occluder on the top (Figure created by authors). (c) A novel, zig-zag variant of White's effect: in this configuration, the structure of T-junctions is inconsistent with either white or black occluder for any of the grey target areas, but, despite this, the induced effect is in the same direction as in the original White's effect (Figure created by authors). (d) The original Wedding cake illusion (Photograph taken by B. Spehar).

Figure III.59-2c depicts a variant of White's stimulus in which the background consists of zig-zag rather than straight stripes. This variant of White's effect was developed by us in 2001 and was designed to eliminate any systematic effects of T-junctions on the difference in perceived lightness between the grey patches lying on the black and white bars. All the target patches still have T-junctions at each corner, but now two of the T-junctions are black along the top and grey/white across the stem and two T-junctions are white across the top and grey/black across the stem. Because of the introduced heterogeneity in the T-junction structure associated with each of the test regions, it is more difficult to argue that they might be receiving any systematic selective contrast integration or inhibition from either white or black inducing regions. Similarly, it is hard to argue for perceptual scission as the explanation of this variant because a consistent decomposition of the test patch luminance into either a lighter or darker transparent region is no longer possible. We have been referring to this effect as either the zig-zag White's effect or the wedding cake illusion, as it featured as our wedding cake design (Fig. III.59-2d). The wedding cake illusion is in the same direction and is of similar magnitude as the classical White's effect (Spehar, Clifford, & Johnston, 2002b).

WHY THE EFFECT IS INTERESTING AND RELEVANT FOR VISION RESEARCH

The wedding cake variant of the original White's effect is one of the most effective demonstrations that rules out the critical contribution of T-junction–based contrast weighting or perceptual scission–based mechanisms. The zig-zag configuration has been designed to balance the effects of T-junctions on the difference in perceived lightness between the grey patches lying on the black and white bars. All the target patches have T-junctions at each corner, two of which are black along the top and grey/white across the stem and two of which are white across the top and grey/black across the stem.

In addition, this variant of the original White's effect is useful in challenging a different class of accounts that focus on the role of early stage inhibitory interactions identical to those thought to be involved in simultaneous contrast effects (Foley & McCourt, 1985; Kingdom & Moulden, 1991; Moulden & Kingdom, 1989). In particular, Blakeslee and McCourt (1999, 2004) have proposed a multiscale spatial filtering model of a range of phenomena in lightness perception including White's effect. Their model combines the responses of a family of oriented cortical filters to produce quantitative estimates of perceived lightness as a function of spatial position within the image. Key to the model's ability to predict the direction of White's effect is the fact that the outputs of the oriented filters are normalized by the pooled activity of all similarly oriented filters. This pooling occurs across the entire image, and, given that the striped background of the White's effect stimulus (Fig. III.59-2a) provides almost all the oriented structure in the image, normalization serves to attenuate the effect of contrast across the gray target's long borders with the horizontal stripes relative to the effect of contrast across the short vertical edges. Since the ratio of horizontal to vertical structure in the whole image is far greater than the aspect ratio of a given test patch, this divisive normalization more than overcomes the effects of adjacency on apparent lightness. Thus the response normalization stage of the model is effectively a computational implementation of the notion that White's effect is due to contrast induced along rather than across the stripes in the stimulus. One perhaps unforeseen consequence of the role of response normalization in Blakeslee and McCourt's (1999, 2004) filtering model is that it predicts that White's effect should be abolished if the ratio of horizontal to vertical structure in the image is equalized, as in the zig-zag variant we created. Even a cursory inspection of Figure III.59-2c suggests that this is not the case.

The zig-zag variant succeeds not only in challenging some of the widespread explanations of the original White's effect but also strongly suggests that assimilation effects seem to prevail in a variety of complex configurations that, based on a range of theoretical approaches to lightness perception, should result in contrast induction.

VARIANTS AND RELEVANT PARAMETERS FOR THE EFFECT

An alternative approach to explaining assimilation phenomena as a complex summation of simultaneous contrast effects is to examine the unique luminance relationships between the inducing elements and the test patch in the inducing configuration. For example, Spehar, Gilchrist, and Arend (1995, 1997) found that, for achromatic stimuli, White's effect occurs only when the luminance of the grey patches lies between the minimum and maximum luminance values of the inducing stripes (see also Spehar, Clifford, & Agostini, 2002). When the luminance of the test regions lies outside the range of the inducing stripes, the test patch surrounded more by regions of lower luminance appears lighter than an identical test patch surrounded more by regions of higher luminance. Figures III.59-3a and III.59-4a show two White's configurations with target regions of highest and lowest luminance, respectively. The difference in the appearance of the two targets is not as salient as in the case of the classical White's effect but is in the opposite direction, consistent with simultaneous contrast. Spehar et al. (1995, 1997) refer to this dependence as the "luminance constraint," while Ripamonti and Gerbino (2001) refer to configurations with "double-incremental" and "double-decremental" targets as the "inverted" White's effect.

Existing models of White's effect focus almost exclusively on the classical version, with target regions of intermediate luminance relative to those of the inducing stripes, and have been largely silent with respect to this dependency. Accounts based on grouping (Ross & Pessoa, 2000), or a T-junction–specific contrast inhibition (Todorović, 1997; Zaidi et al., 1997), would be similarly unsuccessful because the luminance differences between the inverted and classical versions of White's effect are inconsequential for the operation of such mechanisms.

While it is still unclear why the direction of induction changes depending on the relative luminance relationships between the target and inducing regions in White's

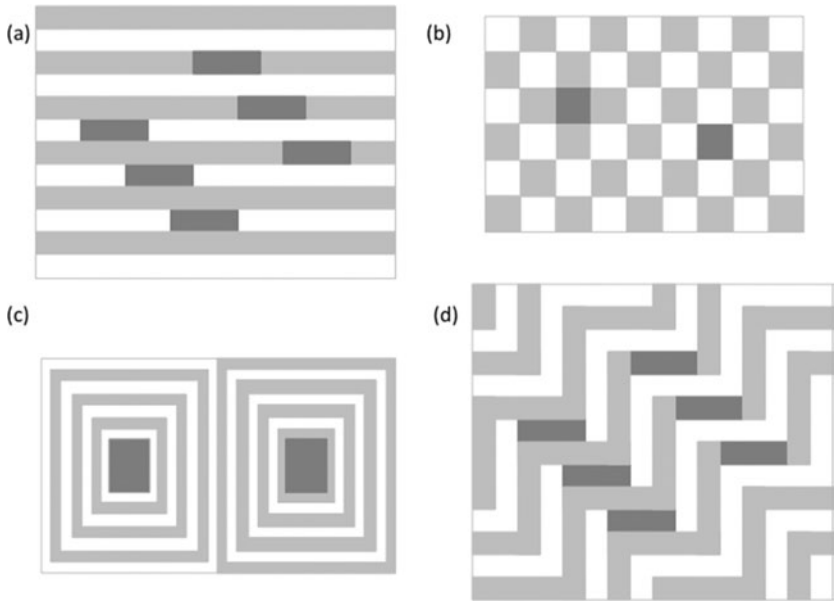


Figure III.59-3. The role of luminance relationships in classical assimilation phenomena. (a) Double-decrement variant of White's effect. (b) Double-decrement variant of the checkerboard illusion. (c) Double-decrement variant of the bullseye illusion. (d) Double-decrement variant of the zig-zag White's illusion. (Figures created by authors.)

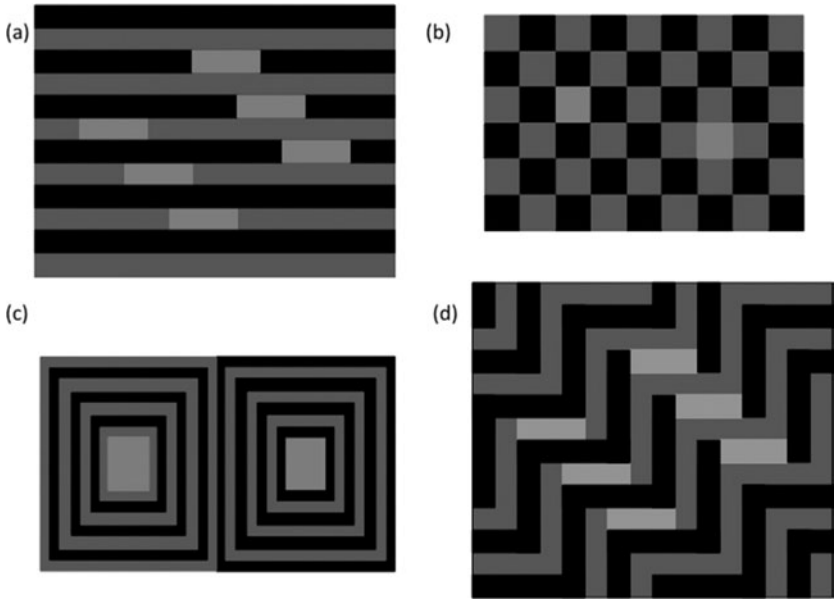


Figure III.59-4. The role of luminance relationships in classical assimilation phenomena. (a) Double-increment variant of White's effect. (b) Double-increment variant of the checkerboard illusion. (c) Double-increment variant of the bullseye illusion. (d) Double-increment variant of the zig-zag White's illusion. (Figures created by authors.)

configuration, it is interesting to point out that this seems to be a very general trend across the well-known assimilation patterns we considered. For example, Figures III.59-3b–d and III.59-4b–d depict double-decrement and double-increment versions of the checkerboard, bullseye, and wedding cake illusory configurations, respectively. In all configurations, the target region surrounded more by lighter-inducing regions appears darker than the target region surrounded more by darker-inducing regions.

Spehar et al. (1995, 1997) have taken these findings as evidence that qualitative relationships between the luminance of inducing and target regions are a crucial determinant of the type of the induction effect observed. Especially in configurations that are geometrically more complex, contrast effects seem to be predominantly restricted to a specific range of photometric relationships: they are observed when the luminance of the target regions is not intermediate to the luminance of the inducing regions. Conversely, when the luminance of target regions is within the range of luminances of the inducing regions, assimilation effects are most typically observed. It is quite possible that additional mechanisms involved in the perception of three-dimensional layout and transparency might be engaged by certain configurations that afford these types of perceptual inference.

Taken together, these observations suggest that induction in the classical and inverted White's and other assimilation configurations is likely due to different underlying mechanisms. While target appearance in the inverted configuration is consistent with lateral inhibition between local, luminance-sensitive mechanisms, accounts of the classical White's and other assimilation effects invariably

involve more global mechanisms. For example, the presumed mechanisms of perceptual scission compatible with illusory transparency operate within the same luminance constraint first reported by Spehar et al. (1995, 1997).

DISCUSSION OF RESEARCH AND THEORIES

Recently several proposals have emerged suggesting that assimilation occurs for a given region when its local contrast is small in comparison to the contrast difference of edges among inducing regions. In this framework, the higher contrast among the inducing edges may partially overcome the lower-contrast signal associated with the local target edge, causing assimilation to occur. Chubb, Sperling, and Solomon (1989) reported that changing only the variance (i.e., variations from the mean luminance level) but not the mean luminance itself in a target's surround induces a profound change in target appearance. This is illustrated in the so-called induced contrast contrast display where two medium-contrast patches are embedded in either a homogenous grey background or a high-contrast texture. Although the mean luminance of the surround in the high-contrast texture is identical to the luminance of the homogenous background, the test patch embedded in the high-contrast surround is perceived as having a lower contrast than the patch embedded in the homogenous surround.

As applied to the phenomenon of assimilation, these models are based on contrast-sensitive neurons responding to the remote contrast of the surround. If these cells are stimulated by a high contrast between surrounding inducers, they can suppress the contrast between the target patch and its adjacent inducers (Bindman & Chubb, 2004; Guclu & Farell, 2005). The high contrast between the remote inducers acts to suppress the contrast of the test patch and its immediate local surround. For example, in the bullseye illusion, the black and white surrounding bands are at almost twice the Weber contrast as compared to the contrast between the test patch and the innermost band (Bindman & Chubb, 2004). This results in a reduced difference in the perceived luminance of the test patch and adjacent band and therefore assimilation. Similarly, in White's illusion the contrast between black and white stripes can suppress the contrast between grey target patches and the abutting inducers. Furthermore, Bindman and Chubb observed that assimilation effects were apparent only when the contrast between a test patch and the average luminance of abutting inducers was small and diminished as it became higher. However, current contrast contrast models do not specify the amount, or even the range, of contrast required *between* inducing elements for assimilation effects to occur.

A similar contrast suppression based contribution is also suggested in edge integration models (Rudd & Arrington, 2001; Vladusich, Lucassen, & Cornelissen, 2006). According to these models, contrast effects induced by a region that is adjacent to a target region (contrast induced at a local edge) can be mediated by contrast effects induced from remote regions (contrast induced from a remote edge). In particular, if the remote edge has a higher luminance contrast than the local edge, it can act to suppress the contrast along the local edge of the target.

CONCLUSIONS

Although assimilation effects tend to occur within more complex contexts, often containing repetitive patterns, we are still generally unable to ascertain the particular conditions that determine whether assimilation or contrast will result. It is generally accepted that as the stimuli increase in complexity, the spatial and photometric characteristics of each surround element become more crucial in determining the central target's appearance. Whether purely local lateral interaction or a subsequent stage of more global image analysis proves to be a more complete account of brightness induction—of both assimilation and contrast varieties—remains to be determined. We anticipate that the effects of geometrical and luminance configurations on the magnitude and direction of induction reported here will be valuable in refining our understanding of the mechanisms of induction phenomena in complex surrounds.

REFERENCES

Agostini, T., & Galmonte, A. (2002). Perceptual organization overcomes the effect of local surround in determining simultaneous lightness contrast. *Psychological Science*, *13*(1), 88–92.

Albers, J. (1975). *Interaction of color.* New Haven, CT: Yale University Press.

Anderson, B. (1997). A theory of illusory lightness and transparency in monocular and binocular images: The role of contour junctions. *Perception*, *26*(4), 419–453.

Anderson, B. L. (2003). Perceptual organization and White's illusion. *Perception*, *32*(3), 269–284.

Bindman, D., & Chubb, C. (2004). Brightness assimilation in bullseye displays. *Vision Research*, *44*(3), 309–319.

Blakeslee, B., & McCourt, M. E. (1999). A multiscale spatial filtering account of the White effect, simultaneous brightness contrast and grating induction. *Vision Research*, *39*(26), 4361–4377. doi:10.1016/S0042-6989(99)00119-4

Blakeslee, B., & McCourt, M. E. (2004). A unified theory of brightness contrast and assimilation incorporating oriented multiscale spatial filtering and contrast normalization. *Vision Research*, *44*(21), 2483–2503.

Chubb, C., Sperling, G., & Solomon, J. A. (1989). Texture interactions determine perceived contrast. *Proceedings of the National Academy of Sciences USA*, *86*(23), 9631–9635.

De Valois, R. L., & De Valois, K. K. (1988). *Spatial vision*. New York: Oxford University Press.

De Weert, C. M. M., & Spillman, L. (1995). Assimilation: Asymmetry between brightness and darkness? *Vision Research, 35*(10), 1413–1419.

Foley, J. M., & McCourt, M. E. (1985). Visual grating induction. *Journal of the Optical Society of America A*, *2*(7), 1220–1230. doi:10.1068/p5348

Gilchrist, A. L., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., . . . Economou, E. (1999). An anchoring theory of lightness perception. *Psychological Review*, *106*(4), 795–834.

Guclu, B., & Farell, B. (2005). Influence of target size and luminance on the White–Todorovic effect. *Vision Research*, *45*(9), 1165–1176.

Heinemann, E. G. (1955). Simultaneous brightness induction as a function of inducing- and test-field luminances. *Journal of Experimental Psychology*, *50*(2), 89–95.

Helson, H. (1963). Studies of anomalous contrast and assimilation. *Journal of the Optical Society of America*, *53*(1), 179–184.

Hering, E. (1964). *Outlines of a theory of the light sense*. Cambridge: Harvard University Press.

Hong, S. W., & Shevell, S. (2004). Brightness contrast and assimilation from patterned inducing backgrounds. *Vision Research*, *44*(1), 35–43.

Kingdom, F. A. A. (2011). Lightness, brightness and transparency: A quarter century of new ideas, captivating demonstrations and unrelenting controversy. *Vision Research*, *51*(7), 652–673.

Kingdom, F. A. A., & Moulden, B. (1991). White effect and assimilation. *Vision Research*, *31*(1), 151–159.

Moulden, B., & Kingdom, F. A. A. (1989) White effect: A dual mechanism. *Vision Research, 29*(9), 1245–1259.

Ripamonti, C., & Gerbino, W. (2001). Classical and inverted White's effects. *Perception*, *30*(4), 467–488.

Ross, W. D., & Pessoa, L. (2000). Lightness from contrast: A selective integration model. *Perception & Psychophysics*, *62*(6), 1160–1181.

Rudd, M. E., & Arrington, K. F. (2001). Darkness filling-in: A neural model of darkness induction. *Vision Research*, *41*(27), 3649–3662.

Shapley, R., & Reid, R. C. (1985). Contrast and assimilation in the perception of brightness. *Proceedings of the National Academy of Sciences USA*, *82*(17), 5983–5986.

Spehar, B., Clifford, C. W. G., & Agostini, T. (2002a). Induction in variants of White's Effect: Common or separate mechanisms? *Perception*, *31*(2), 189–196.

Spehar, B., Clifford, C. W. G., & Johnston, A. (2002b). The role of oriented filters and T-junctions in White's effect. *Journal of Vision*, *2*(7), 555.

Spehar, B., Gilchrist, A. L., & Arend, L. E. (1995). The critical role of relative luminance relations in White's effect and grating induction. *Vision Research*, *35*(18), 2603–2614.

Spehar, B., Gilchrist, A. L., & Arend, L. E. (1997). Qualitative boundaries critical in square-wave brightness illusions: A reply to Kingdom et al. (1996). *Vision Research*, *37*(8), 1044–1048.

Todorović, D. (1997). Lightness and junctions. *Perception*, *26*(4), 379–394.

Vladusich, T., Lucassen, M. P., & Cornelissen, F. W (2006). Edge integration and the perception of brightness and darkness. *Journal of Vision*, *6*(10), 1126–1147.

White, M. (1979). New effect of pattern on perceived lightness. *Perception*, *8*(4), 413–416.

Zaidi, Q., Spehar, B., & Shy, M. (1997). Induced effects of backgrounds and foregrounds in three-dimensional configuration: The role of T-junctions. *Perception*, *26*(4), 395–408.

Chapter 60
Filling-In Between Contours

Rob van Lier

INTRODUCTION

The visual world is a creation of our brain. Contrary to what we may believe, everything we see in fact belongs to the output of the visual process. Starting with meaningless patches of light on our retinae, the visual system rapidly transforms the incoming information into the world we all know so well. Every second our eyes are open, our "mental canvas" is bombarded with colors and forms—a game of light that seems to be played by a hidden artist. Can we unravel all mysterious and secretive tricks that enable vision? There is still a long but exciting way to go. On that scientific journey, visual illusions provide a perfect entrance to learn more about the underlying mechanisms of the visual brain.

Many different kinds of illusions have been discovered already. Illusions with regard to shape or depth may often be experienced as some kind of cognitive misjudgement, whereas properties like brightness or color usually have such a convincing phenomenological perceptual presence that any "misperception" more directly hints at some ingenious deception delivered by our brain. In this chapter, three illusions illustrate how surface properties like color and brightness may interact with perceived contours and shape our daily illusory vision.

THE DOUBLE NEON COLOR ILLUSION

The first illusion is the so-called double neon color illusion (van Lier, 2002). The original neon color spreading effect was discovered and reported by Varin (1971) and van Tuijl (1975); the latter one referred to the color spreading with the term "neon color." A typical stimulus comprises four black three-quarter disks arranged in a square-like fashion such that the contours of indentations are collinear with respect to each other (often referred to as the "Kanizsa square" after its inventor Gaetano Kanizsa). When colored patches are added to complete these three-quarter disks into full disks, the added color may spread throughout the illusory square, revealing the perceptual characteristics of a transparent colored surface.

A remarkable observation is the following: when such illusory colored surfaces cross each other, the area of overlap between these illusory surfaces shows yet another perceptual appearance. This is shown in Figure III.60-1. The figure comprises a grey panel with yellow/white colored disks. The positions of the disks and their coloring is such that the borders of the yellow color patches are perfectly aligned and appear to trigger four illusory yellow-colored bars. These illusory bars seem to cross each other and create an additional illusory percept at the location where they apparently overlap. First, notice that the yellowish bars appear somewhat darker than the grey surroundings. In addition—and crucially—the area of overlap of two yellowish bars appears to be even darker than the illusory bars themselves. This illusory darkening can be enhanced by adding motion to the display; in Video III.60-1a, the yellow patches, and with that the neon-colored bars, make left–right/up–down movements, which enhances the impression of transparency.

Apparently, due to the interaction between two illusory bars, an additional illusory surface is created. Therefore this illusion can perhaps best be classified as a second-order illusion. Note that the direction of the judged lightness can also go in the opposite direction. For example, when the yellow/white patches on a grey panel are replaced by blue/black patches in an otherwise white panel, the bars may appear lighter than the surroundings; in addition, the area of overlap appears lighter than each of the bars separately (see Video III.60-1b; see also van Lier, 2002).

The general perceptual impression at the regions of overlap (Fig. III.60-1b) is one of transparency. In particular, the percept can be classified as a bistable transparency in which either of the illusory bars can be perceived to be positioned in front. In fact this percept complies with an analysis of the perceived luminance difference at X-junctions where the illusory bars intersect (e.g., Anderson, 1997). In a display with real luminance differences, a bistable transparency is characterized by similar luminance polarities in parallel directions at both sides of an X-junction. In the present neon color illusion, all luminance levels at the illusory X-junctions are obviously the same; however, when applying the same idea to the perceived luminance levels, a scheme supporting a bistable transparency emerges. See, for example, Figure III.60-1: the perceived luminance difference at X-junctions where the illusory bars overlap when going from the background to a single neon-colored bar and when going from that neon-colored bar to the area of overlap are in the same direction (i.e., from relatively

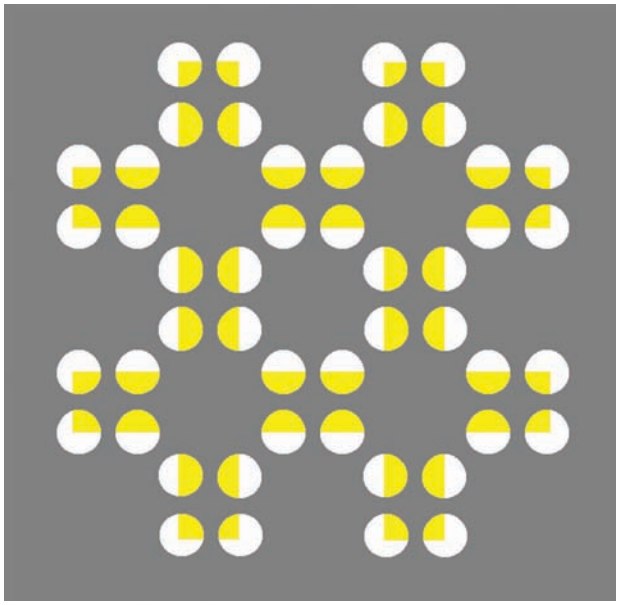
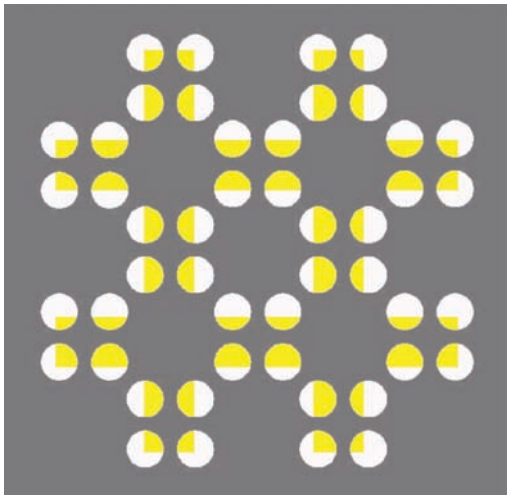


Figure III.60-1. The double neon color illusion. The yellow patches are positioned such that four overlapping bars emerge. Where the illusory, somewhat darker bars cross each other an additional illusory darkening appears. Figure taken from van Lier, R. (2002). A double neon colour illusion. *Perception, 31,* 31–38.

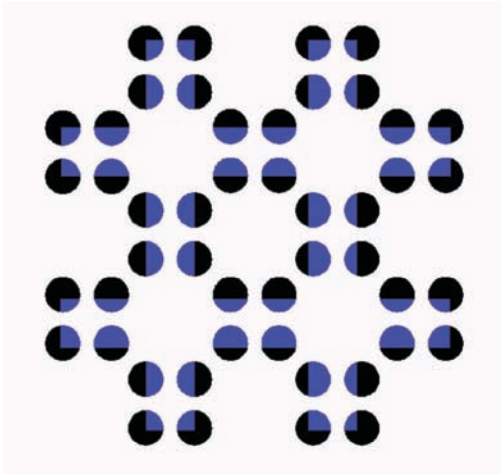
light to a little darker). The double neon color illusion shows that further processing of perceived luminance may reveal a perceptual outcome (transparency) that resembles the percept of a display with similar but real luminance differences.

THE CATCHING PATCHES ILLUSION

This illusion[1] (van Lier & Vergeer, 2006; Vergeer & van Lier, 2010) shows the influence of the mere presence/



Video III.60-1a. The double neon color illusion. The patches induce yellowish illusory bars that appear to be darker than the background. The motion of the bars enhances the perception of the additional darkening at the area of overlap. Animation from van Lier, R. (2002). A double neon colour illusion. *Perception, 31,* 31–38. (The animation should be viewed in the continuous loop mode.)



Video III.60-1b. The patches induce bluish illusory bars, apparently lighter than the background; the motion of the bars further enhances the effect at the area of overlap. Animation from van Lier, R. (2002). A double neon colour illusion. *Perception, 31,* 31–38. (The animation should be viewed in the continuous loop mode.)

absence of luminance contours on the perceptual appearance of adjacent surfaces. When (for example) dark-grey bars on a light-grey background cross each other, the brightness of the area of overlap strongly depends on the presence of outlines around the bars. More specifically, adding outlines may induce somewhat lighter patches at that area. Alternatively, when light-grey bars cross each other on a dark-grey background, additional outlines around the bars would induce somewhat darker patches. In Figure III.60-2, a few examples are shown; in Figures III.60-2a1 and III.60-2b1 we start with a modified Hermann grid, in which (almost) no patches are perceptually present at the crossings.[2] Figure III.60-2a1 shows a grid with dark-grey bars on a light-grey background, whereas Figure III.60-2b1 shows a grid with light-grey bars on a dark-grey background. In each panel, the bars are homogeneously colored (also at the crossings). However, when adding luminance contours, lighter or darker patches emerge at the crossings (see Figures III.60-2a2, III.60-2a3, III.60-2b2, and III.60-2b3).

Two further observations are important here. First, the polarity of the perceived luminance difference at the crossings does not depend on the luminance polarity of the outlines; compare Figures III.60-2a2 and III.60-2a3 where the added contours are relatively dark and light, respectively. Nevertheless, in both cases this results in perceptually lighter patches at the crossings. In a similar way, in Figures III.60-2b2 and III.60-2b3, the dark and light outlines induce perceptually darker patches at the crossings. Apparently the luminance polarity between the grid and the background is crucial; when a grey grid is positioned on a lighter background, the contour-induced patches appear to be lighter, and when a grey grid is positioned on a darker background, the patches appear to be darker (note, however, that the strength of the effect may differ between the displays). Second, the lightening or darkening appears to be captured between the contours—even when the contours have random indentations and protrusions. We reason that without outlines the effects of contrast induction spreads

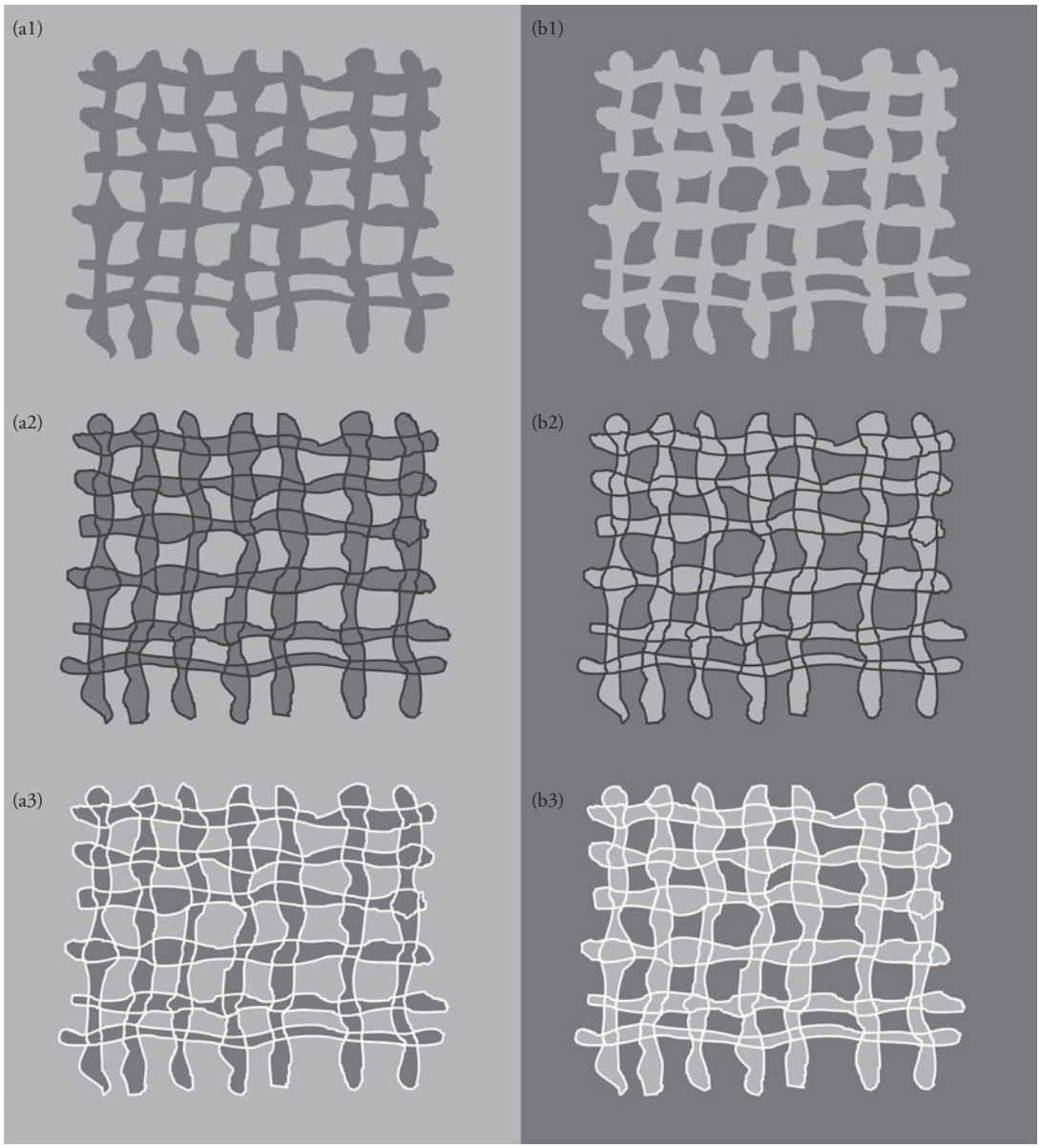
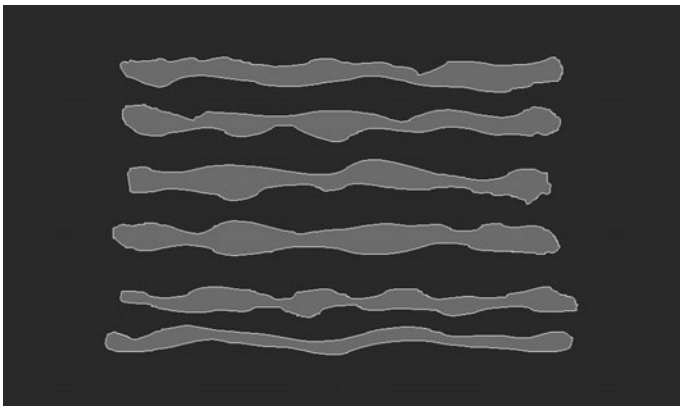


Figure III.60-2. The catching patches illusion. When outlines are drawn around the grey bars, the areas at the crossings appear to have different brightness, irrespective of the luminance polarity of the outlines. Figure adapted from Vergeer, M and van Lier, R. (2010) Capturing lightness between contours, *Perception, 39*, 1565–1578.

throughout the bars, resulting in perceptually homogeneous grey bars and crossings. Adding luminance contours, however, blocks the spreading of the contrast-induced brightness modulation. In Figure III.60-2b1, for example, the grey bars are surrounded by darker regions. When luminance outlines are drawn (Figures III.60-2b2 and III.60-2-b3), the crossing areas are, obviously, juxtaposed to the physically same grey areas at the other side of the outline. The parts of the bars between the crossings, however, are juxtaposed to darker areas (except for the crossing areas) and therefore appear lighter as compared to the crossings. So the catching patches phenomenon seems to be based on simultaneous brightness contrast followed by an additional capturing of the induced brightness by the luminance borders, which prevents a homogeneously spreading throughout the bars. In Video II. 60-2a, a dynamic version of the phenomenon can be viewed, showing that the brightness differences between the crossings and the rest of the bars occur instantaneously. Note that a similar effect may arise when colors are involved. In Video III.60-2b, grey bars overlap each other on a red background. Now the crossings appear to be slightly reddish as compared to the rest of the bars (which may appear somewhat greenish due to color contrast). Video III.60-2c shows a similar example for grey bars on a blue background, inducing slightly bluish patches.

THE AFTERIMAGE FILLING-IN ILLUSION

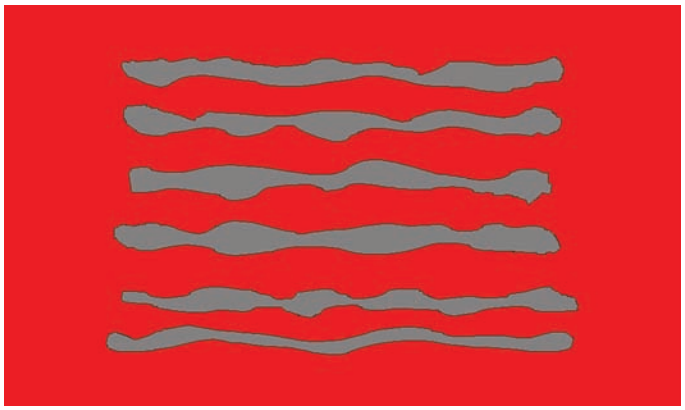
Filling-in phenomena not only occur with "real" colors; they may also occur with afterimage colors. Every student of



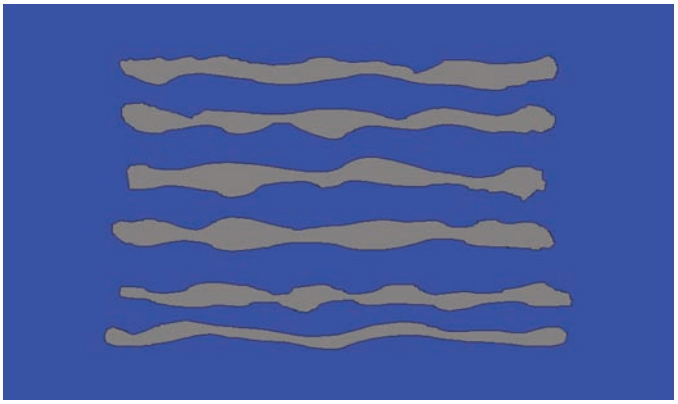
Video III.60-2a. A dynamic example of the catching patches illusion. The moment the grey bars overlap each other, the different lightness at the crossings appear, spread homogeneously between the contours at the crossings. The crossings and the rest of the bars are the same grey. Animation from Vergeer, M. and van Lier, R. (2010) Capturing lightness between contours, *Perception, 39*, 1565–1578.

perception knows that when one adapts to a colored image and then looks at a white surface, a colored afterimage can be seen in the complementary color. For example, looking at a reddish color induces a greenish afterimage and vice versa. Daw (1962) showed that contours may modulate the strength of afterimages. More recently we have shown (van Lier & Vergeer, 2007; van Lier, Vergeer, & Anstis, 2009) that afterimage colors may spread and fill-in between contours presented after the colored image. In addition, the illusion shows that one and the same colored image may induce differently colored afterimages at the same retinal location.[3]

Consider, for example, the eight-pointed colored star in Figure III.60-3; four points are reddish, and the other four points are cyanish. When, in a sequential presentation, this eight-pointed colored star is followed by a four-pointed star outline (referred to as the test outline) that captures the points that previously were red, a cyanish afterimage can be seen not only at the previously colored areas but throughout the whole outlined area. In a similar way, when the colored star is followed by a test outline that captures the previously presented cyanish points, the afterimage in the outlined area appears to be reddish. Apparently



Video III.60-2b. Catching colored patches. The bars are grey; nevertheless, the crossings appear slightly reddish.



Video III.60-2c. Catching colored patches. The bars are grey; the crossings appear slightly bluish.

the colors of the afterimage spread between the contours that are presented *after* the colored stimulus. Moreover, the phenomenon shows that afterimage colors may appear at retinal locations where no color was presented. Readers may wish to consult the animations to see the effect. In Video III.60-3a, the eight-pointed colored star alternates with both test outlines, revealing the differently colored afterimages, and in Video III.60-3b, the eight-pointed colored star is followed by the two test outlines, one immediately after the other; in the latter case the afterimage color rapidly switches from one color to the other (i.e., without an intermediate colored star stimulus). Notice that the spreading effect not only occurs with luminance contours but also shows up when the luminance contours are replaced by illusory contours, although it is somewhat weaker (see Video III.60-3c for an example).

The colored afterimages are not only due to the afterimages of the colored points that are captured inside the

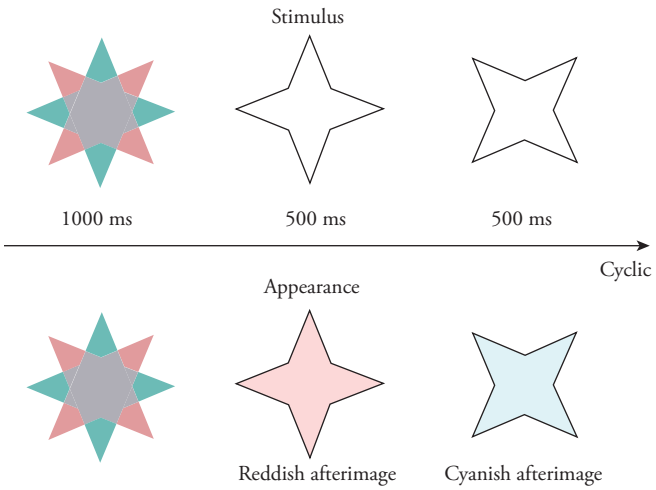
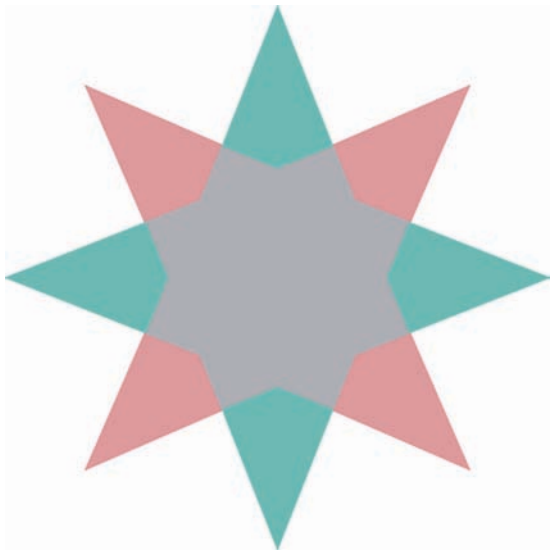


Figure III.60-3. The afterimage filling-in illusion. The top row indicates the actual stimulus. The bottom row indicates the actual percept. When the colored star alternates with the two outlines, two different afterimages can be seen in a sequential fashion. Reprinted (partly adapted) from Current Biology, 19, van Lier, R., Vergeer, M, & Anstis, S. Filling-in afterimage colors between the lines. R232-R324, Copyright (2009), with permission from Elsevier.



Video III.60-3a. The afterimage filling-in illusion. The colored star alternates with two differently oriented outlined stars. The afterimage color inside the star depends on the outlines and appears to be either cyanish or reddish. Taken from Current Biology, 19, van Lier, R., Vergeer, M, & Anstis, S. Filling-in afterimage colors between the lines. R232-R324, Copyright (2009), with permission from Elsevier.

subsequent outline. That is, the colored points outside the test outline also affect the afterimage inside the test outline, resulting in a color similar to those colored points—which is apparently modulated by color contrast induction (e.g., Anstis, Rogers, & Henry, 1978). In van Lier et al. (2009), the relative contribution of the colored points inside and outside the subsequent test outline was further explored. This was done by having various color combinations, taken



Video III.60-3b. The colored star alternates with two star outlines, this time one after the other as in Figure III.60-1. The afterimage in the center of the star changes when the outlines changes. Taken from Current Biology, 19, van Lier, R., Vergeer, M, & Anstis, S. Filling-in afterimage colors between the lines. R232-R324, Copyright (2009), with permission from Elsevier.



Video III.60-3c. The colored star alternates with illusory contours showing that the afterimage filling-in can also be triggered by such contours. Taken from Current Biology, 19, van Lier, R., Vergeer, M, & Anstis, S. Filling-in afterimage colors between the lines. R232-R324, Copyright (2009), with permission from Elsevier.

from the so-called Teufel-Wehrhahn color ring (Teufel & Wehrhahn, 2000; a standardized color set with colors that are nearly isoluminant and are nearly equidistant in color space). We created stars with various color combinations: single-colored stars, stars with complementary colors (i.e., colors at opposite positions in the Teufel-Wehrhahn color ring), and stars comprising colors that are orthogonally positioned in the Teufel-Wehrhahn color ring. In an experiment where participants had to judge the afterimage color inside the test star, all pairwise combinations of these possibilities were used, balanced for position and so on. The results showed that the negative afterimages of the colors inside dominated the percept but that they tended to merge with the relatively weak positive afterimages of the colors that were presented at positions outside the star outline.

With regard to the mechanisms responsible for afterimage filling-in, we suggested that they may resemble those responsible for filling-in due to real colors (such as neon coloring). Francis (2010) recently showed that a computational model that proved its value for filling-in of real colors could also account for afterimage coloring, although a later study showed that it could not cover all aspects (e.g., when the size of the test outline was changed as compared to the size of the inducing colored star stimulus; Kim & Francis, 2011). There is an ongoing discussion about the brain level at which such filling-in might occur. Komatsu (2006) has argued that filling-in may happen relatively early in the visual system, a finding that has been disputed by others (e.g., Cornelissen, Wade, Vladusich, Dougherty, & Wandell, 2006). The afterimage filling-in phenomenon nevertheless suggests that afterimage colors undergo further processing in early visual levels in the brain; see also Shimojo, Kamitani, and Nishida (2001), who reported afterimages of filled-in surfaces. Also, our additional demonstration (van Lier et al., 2009) that afterimage filling-in depends not only on luminance contours but may also be constrained

by, for example, illusory contours further hints at processing in those areas. Other examples can be seen in Anstis, Vergeer, and van Lier (2012) and Hazenberg and van Lier (2013). Indeed, recent brain imaging findings show activation in early visual areas when subjects experience afterimage filling-in (e.g., Hong & Tong, 2011).

Modulations of the filling-in illusions mentioned in this chapter may further reveal properties of the underlying mechanism dealing with color-contour interactions. Perhaps one day our puzzled brains, wondering about all kinds of illusory appearances, may understand their own colorful machinery, knowing that perception is nothing but a wonderful illusion anyway.

NOTES

1. In 2006 this illusion was selected as one of 10 illusions in the yearly "Best Illusion of the Year" event organized by the Neural Correlate Society (van Lier & Vergeer, 2006).
2. Note that when having a similar grid with straight bars, patches could be seen at the crossings (i.e., the so-called Hermann grid); it was shown by Geier, Bernáth, Hudák, and Séra (2008) that modifying the shape of the bars in the Hermann grid may strongly disrupt the occurrence of salient spots at the crossings of the grid (an observation that undermined the textbook explanation of the Hermann grid illusion dealing with the center-surround structure of the receptive fields of retinal ganglion cells).
3. In 2008 this illusion won the first place in the yearly "Best Illusion of the Year" event organized by the Neural Correlate Society (van Lier & Vergeer, 2008).

REFERENCES

Anderson, B. (1997). A theory of illusory lightness and transparency in monocular and binocular images: the role of contour junctions. *Perception*, *26*, 419–454.

Anstis, S., Rogers, B., & Henry, J. (1978). Interactions between simultaneous contrast and coloured afterimages. *Vision Research*, *18*, 899–911.

Anstis, S., Vergeer, M., & van Lier, R. (2012). Luminance contours can gate afterimage colors and "real" colors. *Journal of Vision*, *12*(10), 2, 1–13,

Cornelissen, F. W., Wade, A., Vladusich, T., Dougherty, R., & Wandell, B. (2006). No functional magnetic resonance imaging evidence for brightness and color filling-in in early human visual cortex. *Journal of Neuroscience*, *26*, 3634–3641.

Daw, N. (1962). Why afterimages are not seen under normal circumstances. *Nature*, *196*, 1143–1145.

Francis, G. (2010). Modeling filling-in of afterimages. *Attention, Perception, & Psychophysics*, *72*, 19–22.

Geier, J., Bernáth, L., Hudák, M., & Séra, L. (2008). Straightness as the main factor on the Hermann grid illusion. *Perception*, *37*, 651–665.

Hazenberg S. J., & van Lier, R. (2013). Afterimage watercolors: An exploration of contour-based afterimage filling-in. *Frontiers in Psychology*, *4*, 707. doi:10.3389/fpsyg.2013.00707

Hong, S., & Tong, F. (2011). Neural Representation of form-based color filling-in in early visual cortex [Abstract]. *Journal of Vision, 11*, 379.

Kim, J., & Francis, G. (2011). Color selection, color capture and afterimage filling-in. *Journal of Vision*, *11*, 1–18.

Komatsu, H. (2006). The neural mechanisms of perceptual filling-in. *Nature Reviews Neuroscience*, *7*, 220–231.

Shimojo, S., Kamitani, Y., & Nishida, S. (2001). Afterimage of perceptually filled-in surface. *Science*, *293*, 1677–1680.

Teufel, H., & Wehrhahn, C. (2000). Evidence for the contribution of S cones to the detection of flicker brightness and red–green. *Journal of the Optical Society of America A*, *17*, 994–1006.

van Lier, R. (2002). A double neon colour illusion. *Perception*, *31*, 31–38.

van Lier, R., & Vergeer, M. (2006). Catching patches: Filling-in between edges. *Perception*, *35*(ECVP Abstract Suppl.), 47.

van Lier, R., & Vergeer, M. L. T. (2007). Filling in the afterimage after the image. *Perception*, *36*(ECVP Abstract Suppl.), 200–201.

van Lier, R., Vergeer, M., & Anstis, S. (2009). Filling-in afterimage colors between the lines. *Current Biology*, *19*(8), R323–R324.

van Tuijl, H. (1975). A new visual illusion: Neonlike color spreading and complementary color induction between subjective contours. *Acta Psychologica*, *39*, 441–44.

Varin, D. (1971). Fenomeni di contrasto e diffusione cromatica nell'organizzazione spaziale del campo percettivo. *Rivista di Psicologia*, *65*, 101–12.

Vergeer, M., & van Lier, R. (2010). Capturing lightness between contours. *Perception*, *39*(12), 1565–1578.

Chapter 61
The Glare Effect

Daniele Zavagno and Olga Daneyko

INTRODUCTION

In its achromatic version (Fig. III.61-1a–b, e; Video III.61-1), the glare effect consists of the brightness enhancement of a target region surrounded by luminance ramps (or gradients) ranging from black to white. When the background is white and the target region shares its luminance, naïve and expert observers report the target region as self-luminous surrounded by a misty cloud of light. The illusion has been studied mostly using cross configurations, as these show the role played by background luminance: the illusion, in fact, is sensitive not only to the luminance range of the linear ramps but also to the luminance ratios between the target region and background (Zavagno & Caputo, 2001, 2005).

The illusion is rather robust and works well both on paper and on computer screens, both in static and in dynamic displays. In dynamic displays, the ramps increase and decrease in luminance (Zavagno & Bressanelli, 2008; see Video III.61-1), delivering the impression of a dimmer controlling the intensity of the self-luminous region. One way to eliminate the effect is to delimit the target area with an edge: while an impression of luminous mist may still persist on the ramps (and on the background adjacent to

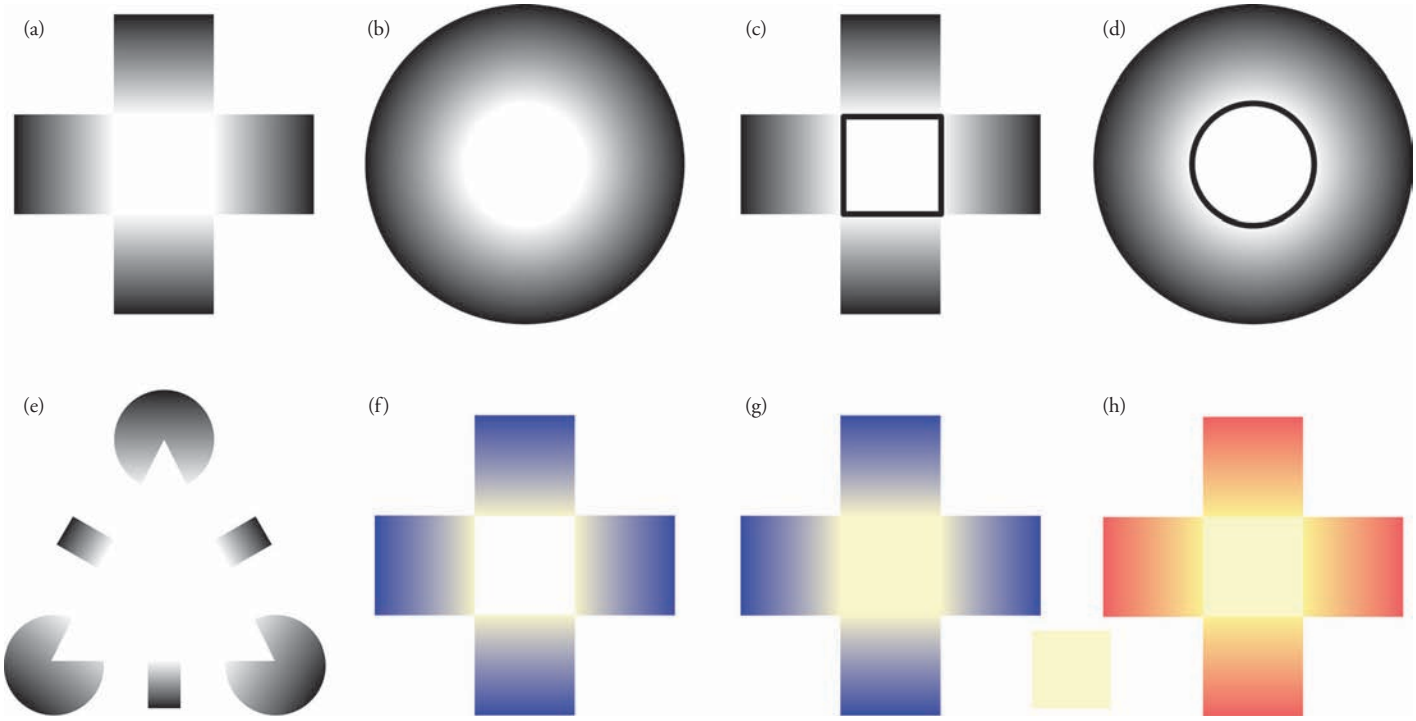
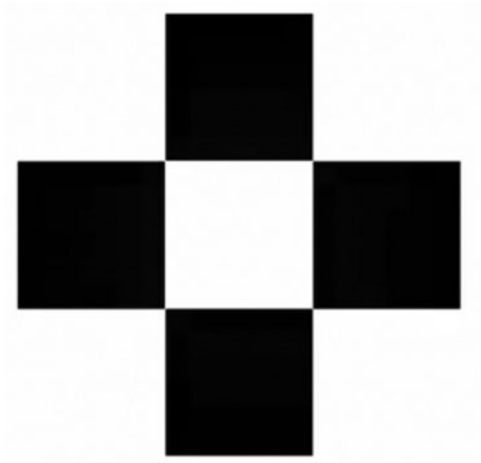


Figure III.61-1. Top row: (a-b), the glare effect in a cross and in a circular configuration (adapted from Zavagno, 1999); (c-d), the brightness enhancement disappears when the target region is delimited by a solid edge. Bottom row: (e), a glowing Kanizsa triangle (adapted from Zavagno, 1999); (f-h), chromatic versions of the glare effect. In (f) the central square appears yellowish while it is actually white, like the background. In (g-h), the central squares are the same color as the smaller square between the two configurations.



Video III.61-1. The dynamic glare effect. The only physical parameter that is changing is the luminance range of the ramp, which from black outward to black inward gradually changes into black outward to white inward.

the cross configuration), the target region appears like an opaque white surface (Fig. III.61-1c–d).

RELEVANCE FOR VISION RESEARCH

The glare effect has been used to study the perception of luminosity (Leonards, Troscianko, Lazeyras, & Ibanez, 2005; Lu, Zavagno, & Liu, 2006; Zavagno & Caputo, 2001, 2005), as it challenges both common sense and scientific theories about luminosity perception. In general, luminosity is considered to be related to an intensity threshold (Ullman, 1976), that is, a luminance value at which point a surface appears to glow. This glowing point probably coincides with what Evans (1974) defined as *fluorence*. In the early days of psychology of perception, it was proposed that a surface appears luminous when it emits or reflects more electromagnetic energy than a surface that appears white within the same visual scene (Hering, 1879; Katz, 1935). This hypothesis implicitly assumes the existence of a luminosity threshold related to the luminance value of a surface that appears white under the same conditions of illumination. Such a *ratio-to-white* hypothesis links inevitably the luminosity threshold to the perception of lightness (i.e., achromatic surface color; Bonato & Cataliotti, 2000; Bonato & Gilchrist, 1994, 1999).

Figure III.61-2 shows the "glare parade,"[1] a series of cross configurations in which the only physical parameter that changes from one cross to the other is the luminance range of the gradients: as the range increases, the brightness of the central square target is also enhanced, from opaque white in Figure III.61-2a, to what may be defined as *superwhite*[2] in Figure III.61-2c–d, to a glowing stage in Figure III.61-2e–f, to an increasing impression of luminosity in Figure III.61-2g–h. Figure III.61 not only challenges the ratio-to-white hypothesis, but it also demonstrates the crucial role played by luminance gradients in luminosity perception over and above just physical intensity (Zavagno, Annan, & Caputo, 2004; Zavagno & Caputo, 2005).

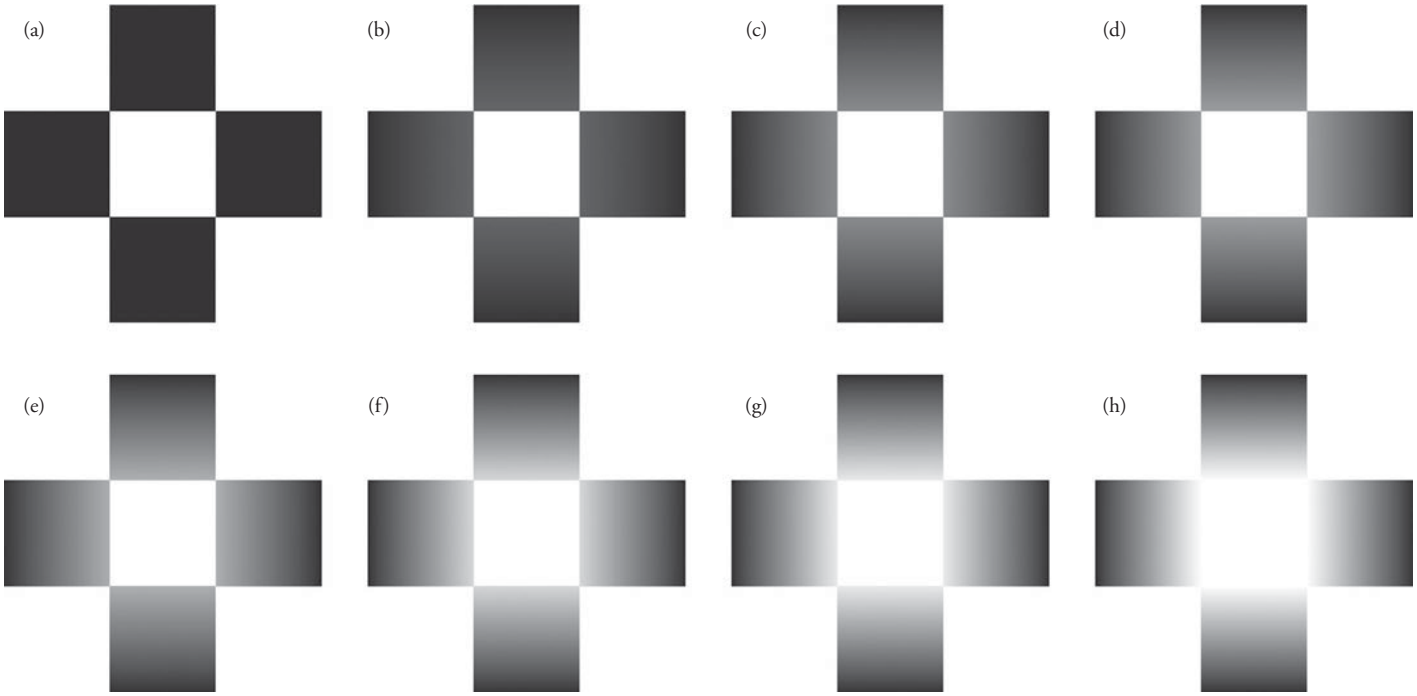


Figure III.61-2. Glare parade: The luminance range of the ramps (gradients) is the only physical parameter that changes from (a) to (h), while the luminance value of the central square targets is equal in all configurations. (Adapted from Zavagno & Caputo, 2005.)

Keil (2007) proposed a neuronal model of the glare effect showing how luminance gradients can be represented by the visual system, suggesting that luminosity perception is in fact related to the presence (and we would also add to the photogeometric distribution) of such gradients. In addition, an empirical study by Correani, Scott-Samuel, and Leonards (2006), aimed at testing luminosity as a visual feature, showed that not luminosity but luminance gradients act as a visual feature. This finding is consistent with the fact that luminance gradients play a crucial role in other brightness-related phenomena, such as the phantom illumination illusion (Zavagno, 2005), which is described later, or shape from shading, in which information for surface curvature, color, and illumination are determined by the distribution of luminance gradients and their relation to other surrounding luminance values (Arnheim, 1974; Zavagno, Daneyko, & Sakurai, 2011).

While the glare effect strongly supports the hypothesis of an independency of luminosity perception from lightness phenomena, there are empirical studies showing how the glare effect actually affects lightness. For instance, in adjustment tasks with glare effect configurations, in which observers were asked to adjust two targets, one to white and the other to luminosity, it was observed that the threshold for white is greater than that for the glare effect and that the presence of the glare effect induces in observers even higher thresholds for white (Zavagno & Caputo, 2005).

The glare effect has been employed also to study the role of luminance gradients on lightness in modified simultaneous lightness contrast (SLC) displays (Agostini & Galmonte, 2002; Zavagno & Daneyko, 2012). When targets are inserted at the center of glowing (Fig. III.61-1) and enhanced darkness squares[3] (Fig. III.61-3a), the global magnitude of SLC is almost doubled (Fig. III.61-5) even if the targets are not adjacent to the luminance ramps. When the direction of the gradients is inverted, so that neither glowing nor enhanced darkness squares are perceived, the magnitude of SLC is practically halved with respect to a standard SLC display (Zavagno & Daneyko, 2012).

VARIANTS AND PARAMETERS

In 1976 Kennedy published a configuration called the *sun figure*, in which a glare effect is obtained by means of dithering. His figure is the first glare effect configuration used for scientific purposes. Nevertheless, the possibility to perceptually enhance the brightness of a certain pictorial region by using colored luminance gradients was employed by European artists way before the official birth of experimental psychology (Zavagno, Daneyko, & Caputo, 2010), for example by Matthias Grünewald (1470/80–1520, Fig. III.61-4), Jacopo Tintoretto (1518–1594), and Rembrandt (1606–1669). By observing their works, one immediately realizes that there are practically infinite combinations of gradients and colors that can determine strong brightness enhancement effects (see Fig. III.61-1f–h).

While the photometric negative of the achromatic glare effect (Fig. III.61-1) determines a darkness enhancement of the central square (Fig. III.61-3a), similar gradients organized in different ways on a dark background can also induce an illumination illusion (Fig. III.61-3b). This last illusion is called the *phantom illumination* (Zavagno, 2005), and it is intrinsically different from the glare effect and its photometric negative in two relevant ways. First, the target area does not appear to glow or be darker but rather brighter than the background, as if it received more illumination. Second, the effect is particularly sensitive to the photogeometric distribution of luminance gradients, as it was found that the phantom illumination is perceived only when the target area appears to belong to the background. The effect is in fact inhibited by the presence of contours, actual or just illusory, delimiting the target region; whereas the brightness of the target region in Figure III.61-3b is modified toward assimilation, in Figure III.61-3c–d it changes toward contrast (Zavagno & Daneyko, 2008b). The glare effect, instead, relies on the presence of contours, physically present, as in Figure III.61-1a–b, or just illusory, as in Figures III.61-1e and III.61-3e.

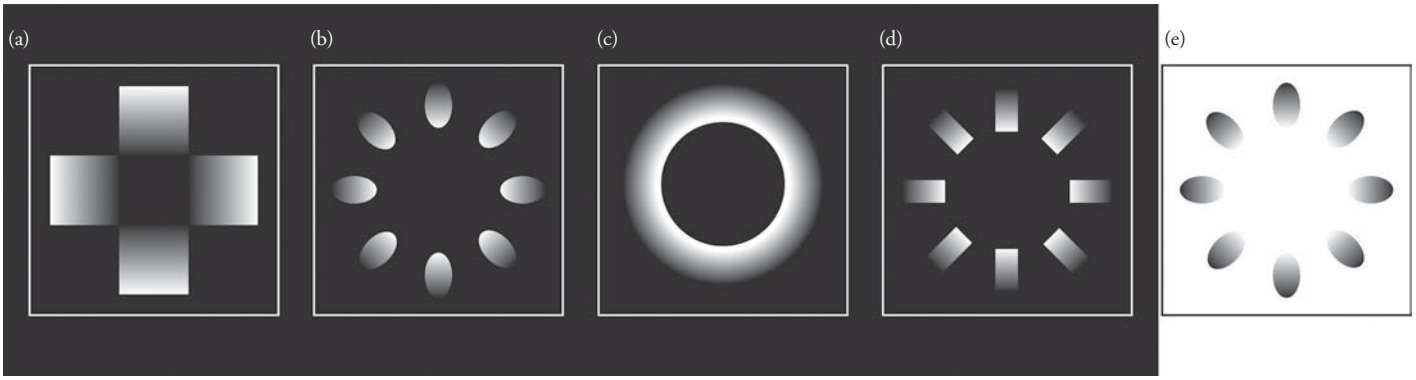


Figure III.61-3. (a), The photometric negative of the glare effect shows a darkness enhancement along with a subtle shape from shading curvature (Zavagno, 1999); (b), the phantom illumination illusion (adapted from Zavagno, 2005); (c-d), actual or illusory contours inhibit the appearance of the phantom illumination illusion (however, illumination appears to be coming from behind the anomalous disk in (d); adapted from Zavagno & Daneyko, 2008); (e), configuration (b) placed on a white background determines a glare effect: notice the presence of an illusory contour in (e) but not in (b).



Figure III.61-4. Matthias Grünewald, *Concerts of Angels and Nativity* (Isenheim Altarpiece, 1515), Colmar: Musée d'Unterlinden.

Finally, an interesting variation of the glare effect can be obtained with so-called luminance pseudo-ramps (Zavagno, 1999; Zavagno & Daneyko, 2008a, Fig. III.61-5a; see also Video III.61-2), in which luminance ramps are replaced by four equivalent series of nonadjacent achromatic rectangles of different luminance, organized within each series from the darkest to the lightest. The resulting configuration shows a brightness enhancement of the central target quite similar to that observable in the glare effect with actual ramps (Fig. III.61-1). Also, these configurations have been used in combination with SLC, producing contrast effects of the same magnitude as those found with actual luminance ramps (Zavagno & Daneyko, 2008a; Fig. III.61-5). The parametric studies inspired by this last type of configuration have just begun; the intention is to revert the problem, using lightness as a tool to study brightness phenomena.

NOTES

1. The name "glare parade" for Figure III.61-2 was suggested some years ago by Ute Leonards.
2. The term "superwhite" is usually used to describe the appearance of a surface that appears brighter than white but is yet not actually self-luminous (Gilchrist, 2006). According to

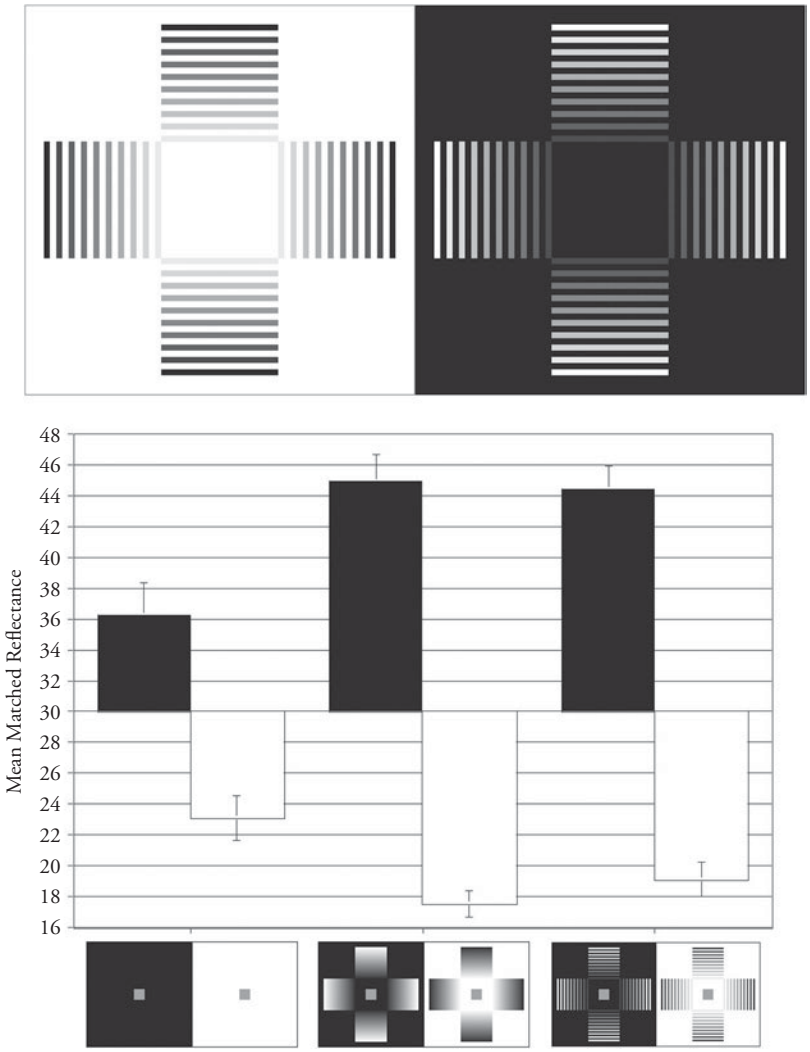
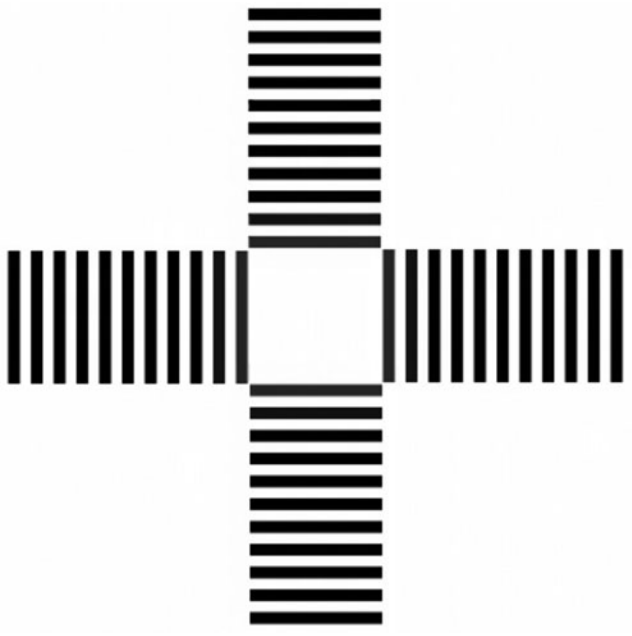


Figure III.61-5. Top: Brightness effects with luminance pseudo ramps. Bottom: Matching results for three types of SLC displays, from left to right: classic display, linear luminance ramps, and pseudo luminance ramps. (From Zavagno & Daneyko, 2008b.)



Video III.61-2. *Neon*, the dynamic version of the glare effect with so-called luminance pseudo-ramps.

Gilchrist, the term is equivalent to Evan's (1974) "fluorence." Such correspondence is, however, questionable, given that Evan's fluorence is closer to the impression of glow and applies also to color (not only to white). Bressan (2006) instead uses the term "superwhite" to indicate a particular condition in which a surface is assigned a temporary lightness value greater than white during one of the stages hypothesized by the double anchoring theory.

3. In the photometric negative of the glare effect, the black square does not appear to glow, but it appears like a penetrable dark space. Zavagno and Caputo (2005) found instead the possibility of generating impressions of luminous grays by employing luminance gradients. Luminous grays have been previously reported by Wallach (1976), yet their existence is somewhat controversial. For instance, Evans (1974) claims that "Grayness and fluorence are mutually exclusive perceptions from a single stimulus" (p. 98), and Wittgenstein (2007) asserts that "Whatever looks luminous does not look grey." (p. 47e)

REFERENCES

Agostini, T., & Galmonte, A. (2002). A new effect of luminance gradient on achromatic simultaneous contrast. *Psychonomic Bulletin and Review, 9*(2), 264–269.

Arnheim, R. (1974). *Art and visual perception: A psychology of the creative eye.* Berkeley: University of California Press.

Bonato, F., & Cataliotti, J. (2000). The effects of figure/ground, perceived area, and target saliency on the luminosity threshold. *Perception & Psychophysics, 62*(2), 341–349.

Bonato, F., & Gilchrist, A. (1994). The perception of luminosity on different backgrounds and in different illuminations. *Perception, 23*(9), 991–1006.

Bonato, F., & Gilchrist, A. (1999). Perceived area and the luminosity threshold. *Perception & Psychophysics, 61*(5), 786–797.

Bressan, P. (2006). The place of white in a world of grays: A double-anchoring theory of lightness perception. *Psychological Review*, *113*(3), 526–553.

Correani, A., Scott-Samuel, N. E., & Leonards, U. (2006). Luminosity: A perceptual feature of light-emitting objects? *Vision Research*, *46*(22), 3915–3925.

Evans, R.M. (1974). *The perception of color.* New York: Wiley.

Gilchrist, A. (2006). *Seeing black and white*. New York: Oxford University Press.

Hering, E. (1879). Der Raumsinn und die Bewegung des Auges. In L Hermann (Ed.), *Handbuch der physiologi*e (Vol. 3, part 1, pp. 341–601). Leipzig, Germany: Vogel.

Katz, D. (1935). *The world of colour.* London: Kegan Paul.

Keil, M. S. (2007). Gradient representations and the perception of luminosity. *Vision Research*, *47*(27), 3360–3372.

Kennedy, J. M. (1976). Sun figure: An illusory diffuse contour resulting from an arrangement of dots. *Perception, 5*(4), 479–481.

Leonards, U., Troscianko, T., Lazeyras, F., & Ibanez, V. (2005). Cortical distinction between the neural encoding of objects that appear to glow, and those that do not. *Cognitive Brain Research, 24*(1), 173–176.

Lu, H., Zavagno, D., & Liu, Z. (2006). The glare effect does not give rise to a longer lasting afterimage. *Perception*, *35*(5), 701–707.

Ullman, S. (1976). On visual detection of light sources. *Biological Cybernetics, 21*(4), 205–212.

Wallach, H. (1976). *On perception*. New York: Quadrangle/ New York Times Book Co.

Wittgenstein, L. (1977). *Remarks on colour.* Berkeley and Los Angeles: University of California Press.

Zavagno, D. (1999). Some new luminance-gradient effects. *Perception, 28*(7), 835–838.

Zavagno, D. (2005). The phantom illumination illusion. *Perception & Psychophysics, 67*(2), 209–218.

Zavagno, D., Annan, V., & Caputo, G. (2004). The problem of being white: Testing the highest luminance rule. *Vision, 16*(3), 149–159.

Zavagno, D., & Bressanelli, D. (2008). The animated glare effect: Another case of perceptual causality? *Teorie e Modelli, 13*(2–3), 313–321.

Zavagno, D., & Caputo, G. (2001). The glare effect and the perception of luminosity. *Perception, 30*(2), 209–222.

Zavagno, D., & Caputo, G. (2005). Glowing grays and surface white: The photo-geometric factors of luminosity perception. *Perception, 34*(3), 261–274.

Zavagno, D., & Daneyko, O. (2008a). The effect of luminance pseudo-ramps on simultaneous lightness contrast. *Perception*, *37*(Suppl.), 147–148.

Zavagno, D., & Daneyko, O. (2008b). When figure-ground segmentation modulates brightness: The case of phantom illumination. *Acta Psychologica, 129*(1), 166–174.

Zavagno, D., & Daneyko, O. (2012). The effect of non-adjacent luminance gradients on simultaneous lightness contrast. In C. Leth-Steensen & R. Schoenherr (Eds.), *Proceedings of the 28th annual meeting of the International Society for Psychophysics* (pp.48–53). Ottawa: International Society for Psychophysics.

Zavagno, D., Daneyko, O., & Caputo, G. (2010). Et facta est lux: Experiencing luminosity in art and in the real world. In *Proceedings of the Third International Workshop on Kansei* (pp. 12–15). Fukuoka, Japan: Across Fukuoka.

Zavagno, D., Daneyko, O., & Sakurai, K. (2011). What can pictorial artifacts teach us about light and lightness? *Japanese Psychological Research*, *53*(4), 448–462.

PART IV
Motion Based

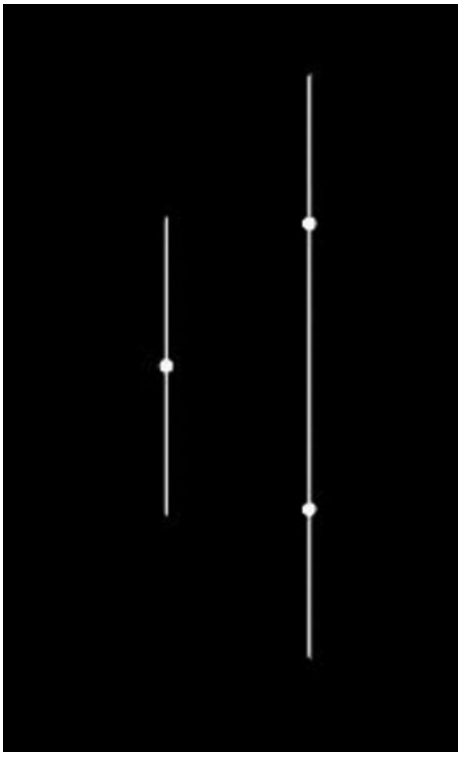
Chapter 62
Improbable Illusory Contours

Barton L. Anderson

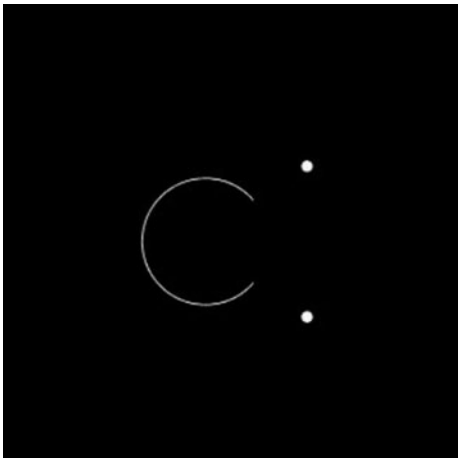
Illusory contours (ICs) are one of the most widely studied kinds of visual illusion. They provide a salient example of the constructive nature of visual processing, in which the brain generates percepts of object boundaries in image regions devoid of any information about the boundary's presence. The need for this constructive process arises from the intermittent camouflage of portions of object boundaries by the coincidental blending of a foreground object with its background. Early debates about the origin of ICs focused on whether they were properly construed as constructed by relatively low-level visual processes or were the outputs of high-level cognitive inference. Over the past few decades, a large body of research has provided evidence that sensitivity to ICs is expressed in a variety of neural loci, including early cortical areas such as V2 (Peterhans, von der Heydt, & Baumgartner, 1984, 1986; Sugita, 1999; von der Heydt & Peterhans, 1989; von der Heydt, Peterhans, & Baumgartner, 1984), as well as higher levels along the visual pathway. It remains unclear whether the IC sensitivity exhibited in early visual pathways is primarily driven by feedback or through the functional connectivity within early visual areas, but these findings have placed IC phenomena within the domain of basic visual processes and inspired a host of psychophysical, neurophysiological, and computational studies into how and when they are generated.

Early studies of IC formation focused on contours formed by static, two-dimensional, monocular images. They were divided into two broad classes: those generated by pairs of oriented contour segments that could be linked into a common path and those generated by the abrupt termination of thin contours or line endings. The former class has attracted the bulk of empirical and theoretical work, presumably because they appear more richly specified and potentially computable by early cortical circuits sensitive to local contour orientation. The latter class of ICs is less constrained because there is no information at the locations where the IC appears to specify its orientation. One approach has been to treat abrupt line endings (or terminators) as potential cues to occlusion, which—based on a variety of psychophysical results—would exhibit a bias to form contours approximately perpendicular to the inducing contour's orientation (Grossberg & Mingolla, 1985a, 1985b). Although such ideas provided some leverage into understanding a variety of IC phenomena, a number of counterexamples to the "perpendicular bias" emerged, most notably in the work of Gillam (1987). Such ideas received even less support when considering the problems of IC synthesis in domains involving multiple views of a scene, such as occurs in stereoscopic and motion displays (Anderson, 1994; Anderson & Barth, 1999; Anderson & Julesz, 1995; Anderson, O'Vari, & Barth, 2011).

The spatiotemporal geometry of occlusion generates features that are either unpaired in space (stereopsis; see Anderson, 1994; Anderson & Julesz, 1995; Gillam & Borsting, 1988; Julesz, 1964; Lawson & Gulick, 1967; Nakayama & Shimojo, 1990) or unpaired in space-time (motion; see Anderson & Sinha, 1997). In motion, these spatiotemporally unpaired features have been termed "deletions" (disappearances) and "accretions" (appearances), which occur as a near surface progressively occludes or uncovers a more distant surface (respectively). In my early work into stereoscopic vision, I discovered a simple but previously unobserved fact about binocular occlusion geometry, which led to the discovery of a new form of quantitative depth and orientation information (Anderson, 1994; Anderson & Julesz, 1995; Malik, Anderson, & Charowhas, 1999). Inspired by the relationship between binocular and motion parallax, I sought to determine if similar phenomena could be observed in the motion domain. Together with Pawan Sinha, we observed a motion analogue of the ICs in stereopsis (Anderson & Sinha, 1997; see Video IV.62-1). The fundamental insight was that the perception of ICs was intimately coupled to the visual system treating certain types of terminator motion as *non-motions* or, more precisely, as forms of accretions and/or deletions. The concept of "motion" refers to the space-time displacement of a particular feature. However, when a contour disappears behind an occluder, the occluded portion of the contour has no correspondence with anything visible and hence does not generate a motion. Likewise, when a portion of a contour (or surface) first becomes visible from behind a moving surface, there is no previous space-time position that defines its motion. We argued that the decomposition of moving contour into motions (persisting regions) and non-motions (accretions and deletions) plays a causal role in the formation of ICs in moving displays.



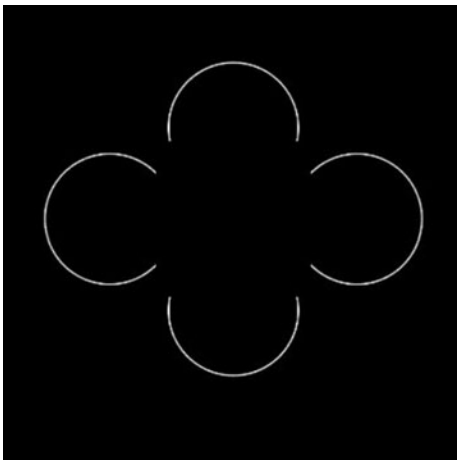
Video IV.62-1. This video demonstrates the importance of accretions and deletions in the genesis of illusory contours. By providing some parallax information suggestive of an occluding surface, the transformations in the lengths of the contours is treated as arising from the appearance (accretion) and disappearance (deletion) of portions of the contours from view, giving rise to a vivid percept of an illusory occluding surface. (From Anderson & Sinha, 1997.)



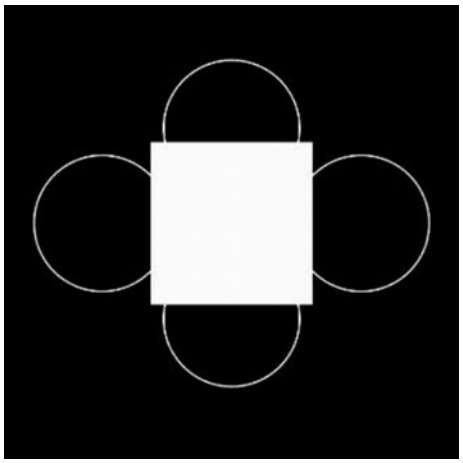
Video IV.62-2. A simulated occlusion sequence where a thin circular contour moves behind a stationary vertical edge. Note the formation of an invisible occluding contour that appears curved and to deform in shape. (From Anderson & Barth, 1999.)

To further assess this possibility, we constructed a simple motion display of a circular outline translating under an invisible occluding surface (Anderson & Barth, 1999). The invisible occluder was a simple, straight, vertical edge, and the circle was translated behind it (see Video IV.62-2). As predicted, we observed a vivid IC generated by the progressive accretion and deletion of the circular contour's boundary. However, much to our surprise, we observed that the illusory-occluding surface was not the same shape as the invisible occluder, and it did not even appear stationary. Rather, the IC appeared to deform in shape as the circular outline was progressively occluded and to move in the opposite direction of the occluded circle. We therefore had two surprising new forms of IC phenomena to explain: a deforming change in shape and a perceived motion. What was the cause of these effects? Our initial attempts focused on treating the perceived motion of the IC as an epiphenomena of mechanisms that generated a specific IC shape in response to the occluded circle's motion. However, we were unsuccessful at constructing any model that could explain the detailed quantitative shape data that we had collected. After many failed attempts, we considered the factor that might have inadvertently inverted the cause–effect relationships responsible for these IC percepts. Perhaps the deformations in perceived *shape* were caused by a pattern of *motion* that was imparted onto the illusory surface, rather than the other way around. The phenomenon of induced motion is well known in the literature and seemed like a plausible candidate. Indeed, we found that the perceived shape of the ICs we observed in these displays could be fully understood if we treated perceived shape as a quantity that is derived from the imparted motion (Anderson & Barth, 1999). This deformation is particularly vivid in Video IV.62-3, which depicts a series of four circular outlines translating behind a stationary, invisible black occluding square (see Video IV.62-4 for the "true" occluding surface). A small phase offset is introduced in the two pairs of opposing disks, causing a percept of a vividly bowing, elastic contour.

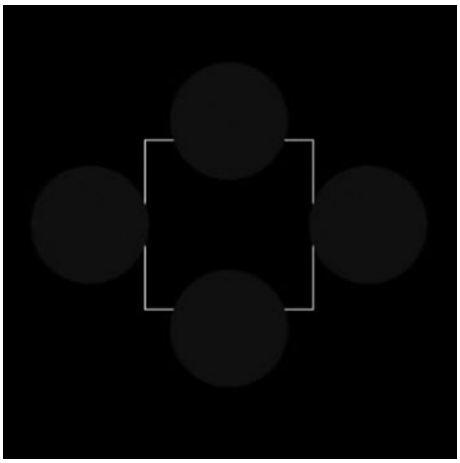
This model, although counterintuitive (and perhaps even wrong), needed to be tested. We began by considering what would happen if the occluded contour, which the model asserts is responsible for the direction and shape imparted onto the IC, was stationary rather than moving. In this case, there would be no way to disambiguate the



Video IV.62-3. A series of four circles are translated behind an invisible stationary occluding square. Note that the shape of the occluding surface does not appear to be square or be stationary, but rather, appears to deform and change shape. Unpublished variant of Anderson and Barth (1999).

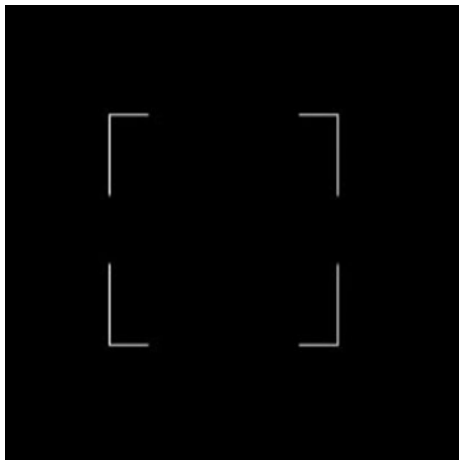


Video IV.62-4. The same sequence as in 62-3, but the occluding surface has now been colored white. (Unpublished variant of Anderson and Barth, 1999.)



Video IV.62-6. The same occlusion sequence as in 62-5, but now the occluding disks have been filled with a different color than the surround. Note that vivid illusory contours are seen inside the outline square, despite the fact that all of the image data is fully accounted for by the visible occluding disks. (From Anderson et al., 2011.)

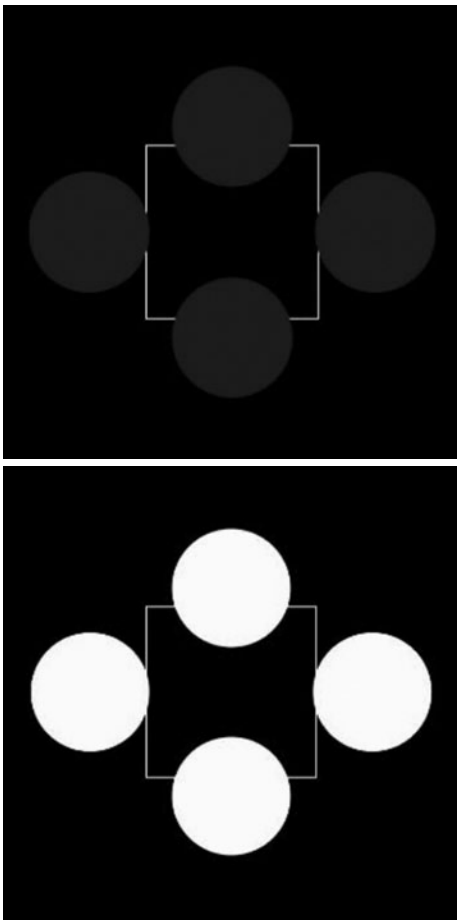
direction of motion of the occluding surface, its shape, or its speed. What ICs would form in response to such a display? To answer this question, we inverted the depth ordering in our display and reversed the role of the occluding and occluded surfaces and contours (Anderson et al., 2011). In this new display, we presented a series of four occluding disks that translated over the sides of a stationary outline square (Video IV.62-5). We began by presenting the occluding disks with the same color as the background. By hypothesis, the pattern of accretion and deletion of the square's edges should elicit a percept of ICs. However, in this case, both the directions of motion of the occluding surfaces and the orientation of the occluding contours are ambiguous: the occluders could originate from outside of the square or from within and have any variety of shapes and velocities. However, if the disks were presented in a different color than the background—rendering the true occluding surface visible—this ambiguity should be eliminated, and no ICs should be observed. Remarkably, however, this is not what happens: vivid ICs appear *within* the occluded square, even though all of the visual data is fully "explained" by the visible occluding disks (Video IV.62-6).



Video IV.62-5. A display in which four disks are translated over the edges of a stationary outline square. Note that illusory contours appear to form over all sides, but the shape and direction of motion of the occluding figures are unclear.

The ICs observed in this display are remarkably robust and exhibit a simple and striking dependence on the relative strength of the different edges in this display. When the contrast of the occluding disks is increased, the strength of the IC slowly weakens and eventually vanishes (Videos IV.62-6, IV.62-7, and IV.62-8). Similarly, if the contrast of the occluded frame is decreased, the IC decreases with it and eventually disappears.

What is responsible for these bizarre and counterintuitive contours? Visual scientists typically approach "explanations" of visual phenomena in one of two ways (which need not be mutually exclusive), neither of which have changed much since they were famously debated between Helmholtz and Hering. One strategy is to construct a mechanistic account that treats percepts as a consequence of some understood neurophysiological property. For example, color afterimages can be explained as the unbalanced response of cones that have been selectively fatigued by lights with different chromatic content. Another explanatory strategy is probabilistic and is usually credited to Al Hazen or Helmholtz. This approach attempts to literally "rationalize" perception by treating perceptual experience as some form of "best guess" about the causes of the input. This is sometimes embedded in an "adaptionist" evolutionary framework, in which particular percepts are treated as forms of probabilistic inference that granted competitive advantages to animals equipped with the best inferential engines. The best of both worlds, in this modeling universe, is to have a mechanistic model that can be identified as performing some computational task optimally, such as the abundant proliferation of Bayesian models of vision (and everything else). At present, it is difficult to understand how probabilistic models could account for the improbable percepts that occur with the IC phenomena described here. The ICs observed in Videos IV.62-6, IV.62-7,



Video IV.62-7. A sequence of movies demonstrating the effect of disk luminance on the strength of the illusory contours. Note at the contrast of the disk relative to the surround gets larger, the strength of the illusory contour gets weaker.

and IV.62-8 entail an extraordinary coincidence of two simultaneous occluding surfaces. In order to demonstrate that they can be rationalized as probabilistically optimal, it would have to be shown that these phenomena "fall out" of more complete analysis of the environmental pressures to generate ICs and, when considered over the ensemble of all possible forms of occlusion, represent a globally optimal solution. At this stage, it is unclear whether such an analysis exists.

Alternatively, and perhaps more plausibly, these improbable ICs represent a mechanistic epiphenomena of circuits used to synthesize ICs, which may not have a simple "rational" explanation that treats them as optimal. At this juncture, it is unknown what causes these ICs to form, which is why they offer a striking opportunity to test theoretical approaches to problems of perception. The ICs described herein are as strong as virtually any that have been reported in the literature and provide some of the most significant challenges to theoretical approaches that treat percepts as forms of optimal inference. The question is whether research will progress under the proverbial streetlight merely because that is where the light is shining, or whether efforts will be focused on phenomena that provide the most difficult challenges to a favored view.

REFERENCES

Anderson, B. L. (1994). The role of partial occlusion in stereopsis. *Nature*, *367*(6461), 365–368.

Anderson, B. L., & Barth, H. C. (1999). Motion based mechanisms of illusory contour synthesis. *Neuron*, *24*, 433–441.

Anderson, B. L., & Julesz, B. (1995). A theoretical analysis of illusory contour formation in stereopsis. *Psychological Review*, *102*, 705–743.

Anderson, B. L., O'Vari, J., & Barth, H. (2011). Non-Bayesian contour synthesis. *Current Biology*, *21*(6), 492–496.

Anderson, B. L., & Sinha, P. (1997). Reciprocal interactions between occlusion and motion computations. *Proceedings of the National Academy of Sciences USA*, *94*(7), 3477–3480.

Gillam, B. (1987). Perceptual grouping and subjective contours. In S. Petry & G. E. Meyer (Eds.), *The perception of illusory contours* (pp. 268–273). New York: Springer-Verlag.

Gillam, B., & Borsting, E. (1988). The role of monocular regions in stereoscopic displays. *Perception*, *17*(5), 603–608.

Grossberg, S., & Mingolla, E. (1985a). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychoogical Review*, *92*(2), 173–211.

Grossberg, S., & Mingolla, E. (1985b). Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations. *Perception & Psychophysics*, *38*(2), 141–171.

Julesz, B. (1964). Binocular depth perception without familiarity cues: Random-dot stereo images with controlled spatial + temporal properties clarify problems in stereopsis. *Science*, *145*(363), 356–362.

Lawson, R. B., & Gulick, W. L. (1967). Stereopsis and anomalous contour. *Vision Research*, *7*(3–4), 271–297.

Malik, J., Anderson, B. L., & Charowhas, C. E. (1999). Stereoscopic occlusion junctions. *Nature Neuroscience*, *2*(9), 840–843.

Nakayama, K., & Shimojo, S. (1990). Davinci stereopsis: Depth and subjective occluding contours from unpaired image points. *Vision Research*, *30*(11), 1811–1825.

Peterhans, E., von der Heydt, R., & Baumgartner, G. (1984). Illusory contour stimuli reveal stages in visual cortical processing. *Perception*, *13*(1), A16–A17.

Peterhans, E., von der Heydt, R., & Baumgartner, G. (1986). Neuronal responses to illusory contour stimuli reveal stages of visual cortical processing. In J. D. Pettigrew, K. J. Sanderson, & W. R. Levick (Eds.), *Visual neuroscience* (pp. 343–351). Cambridge, MA: Cambridge University Press.

Sugita, Y. (1999). Grouping of image fragments in primary visual cortex. *Nature*, *401*(6750), 269–272.

von der Heydt, R., & Peterhans, E. (1989). Mechanisms of contour perception in monkey visual-cortex: I. Lines of pattern discontinuity. *Journal of Neuroscience*, *9*(5), 1731–1748.

von der Heydt, R., Peterhans, E., & Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science*, *224*(4654), 1260–1262.

Chapter 63
Low-Level Motion Illusions

Stuart Anstis

INTRODUCTION

Motion is simply the changes in an object's position over time, so in theory we could perceive motion by sensing the different positions adopted by a moving object over time and dividing the shift by the time taken. But this is not what happens. Instead, motion is a fundamental perceptual dimension, with its own specialized neural mechanisms. Detecting rapidly moving prey or predators in real time is crucial to survival, which is presumably why we have evolved specialized motion detectors; in fact Walls (1942) described motion perception as the most ancient and primitive form of vision. Reichardt (1986) has described a possible "delay-and-multiply" structure for neural motion detectors.

In apparent or stroboscopic motion, an object is flashed up successively in two (or more) different positions. A Reichardt detector will respond to such intermittent stimuli, which can also look convincingly like real movement to human observers. This is the basis of TV and movies. Watson and Ahumada (1985) have modeled human sensitivity to apparent motion. It makes sense in evolutionary terms that we should respond to intermittent, or otherwise poor-quality, stimuli and perceive them as motion, because a false positive, in which we believe we see motion although none is there, is far better than a false negative, in which we fail to notice that a predator is creeping or sprinting toward us. R. L. Gregory has compared motion perception to a lock, which must be designed to accept even ill-fitting keys such as apparent motion (otherwise it would lock out homeowners as well as burglars). Apparent motion has attracted much experimental attention, partly because the stimuli are so tractable and easy to manipulate. This chapter presents various illusions of apparent motion.

CROSSOVER MOTION

In Video IV.63-1, two parallel bars side by side, one dark and one light, switch luminances repetitively over time. (Only the luminances matter; the colors are immaterial.) This generates a stimulus that is consistent with two potential competing bar motions, one dark and the other light. The result is the bar that is seen as moving is whichever bar has the higher contrast, and this depends on the relative luminances of the bars and their surround. On a light surround, the dark bar is seen as moving (Video IV.63-1, Panel 1). On a dark surround, the light bar is seen as moving (Panel 2).

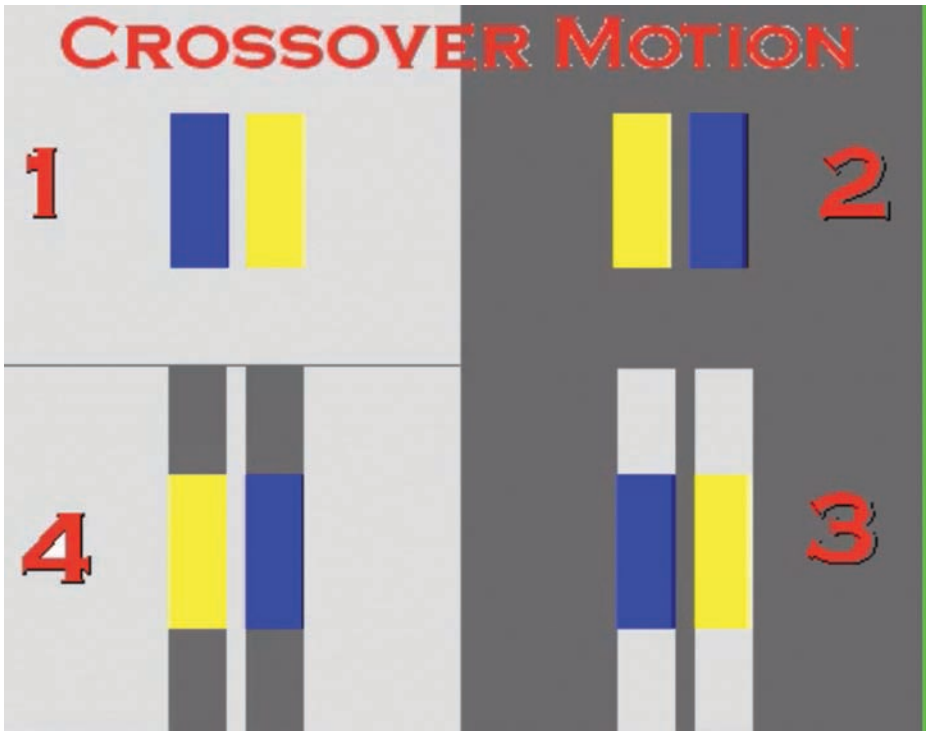
Thus the bar differing more from the surround luminance dominates the motion percept (Anstis, 2003; Anstis & Mather, 1985; Anstis, Smith, & Mather, 2000). This is consistent with the notion of "motion energy," which depends on luminance contrast (Adelson & Bergen, 1985). But this is not the whole story. Although Panels 1 and 4 are the same except for the dark, vertical embedding bars in Panel 4, these are enough to make the light bar move in Panel 4. In other words, motion is computed in the brain after the filtering processes that are responsible for White's effect (Blakeslee, Pasieka, & McCourt, 2005; White, 1981).

REVERSE PHI

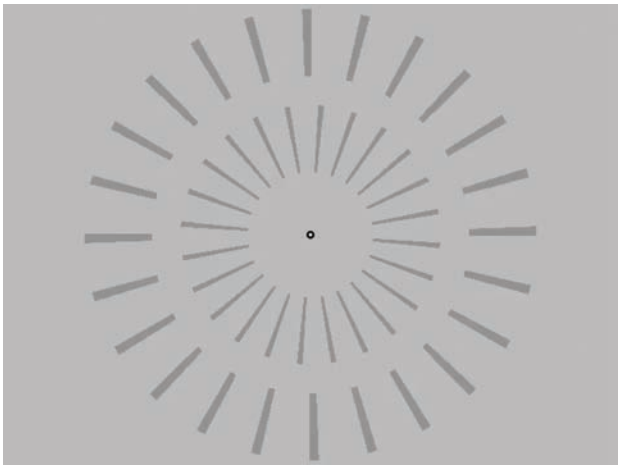
Anstis (1970) and Anstis and Rogers (1975) showed that a two-frame movie of apparent motion appears to go in the reverse direction if the two frames are of opposite contrast. We termed this "reverse phi." Video IV.63-2, kindly supplied by Patrick Cavanagh, shows reverse phi. On fixation of the central spot, the outer ring of radii appear to rotate counterclockwise and the inner ring clockwise. However, inspecting and tracking an individual radius reveals that the outer ring actually rotates clockwise and the inner ring counterclockwise. Reverse phi is consistent with Adelson and Bergen's (1985) motion-energy model.

Wehrhahn (2006) measured the spatiotemporal range over which reverse phi could be seen. He presented pairs of lines on a gray background. Either both were bright or both were dark (equal contrast polarity: normal phi), or one line was bright and the other was dark (opposite contrast polarity: reverse phi). Observers were instructed to indicate the perceived direction of motion. With foveal viewing, reverse phi was seen for small spatial separations (0 to 12 arcmin) and small temporal offsets (8 to 33 ms).

Bours, Kroes, and Lankheet (2009) noted that low-level contrast information in the primary visual pathway is represented in two different channels. On-center cells signal positive contrasts (light spots) and off-center cells signal negative contrasts (dark spots). They quantitatively compared motion coherence detection for regular and for



Video IV.63-1. A light and a dark bar exchange places. The bar with the higher contrast against the background is the one that appears to move—the dark blue bar on the light surround in (1), the light yellow bar on the dark surround in (2). When the bars are embedded in stationary stripes, results are consistent with White's effect—the light yellow bar embedded in a dark stripe appears to move in (4), and the dark blue bar embedded in a light stripe appears to move in (3). (Video by Stuart Anstis.)



Video IV.63-2. In this video, made by Patrick Cavanagh, the radial spokes reverse their luminance polarity on every frame. This makes them appear to move opposite to their physical displacement. During fixation on the central spot, the outer radii appear to rotate clockwise and the inner radii counterclockwise. If the motion is then stopped, the outer radii show a counterclockwise aftereffect of motion and the inner ones show a clockwise aftereffect. But if we look directly at an outer radius, we can see that it is actually moving around counterclockwise. Reverse phi is consistent with Adelson and Bergen's (1985) model of motion energy. (Video by Patrick Cavanagh, Dartmouth U.)

reverse phi motion stimuli. In reverse phi motion, the contrast of a pattern flips during displacements, so sensitivity is therefore based on correlating positive and negative contrasts, whereas for regular motion it is based on correlating similar contrasts. They concluded that reverse phi is perceived through an efficient combination of signals from on and off cells. The same authors (Bours, Kroes, & Lankheet, 2007) compared reverse phi to the motion aftereffect, arguing that motion adaptation causes reduced activity during a stationary test stimulus, which by means of directional opponency leads to motion perceived in the opposite direction. Their results led them to suggest that reverse phi motion similarly reduces the activity of low-level motion detectors.

Livingstone and Conway (2002) examined responses to phi and reverse phi apparent motion stimuli in 118 V1 cells in alert macaques. *All* of the cells showed direction-selective responses to two-bar apparent-motion stimuli, and all of them showed reversed direction preference when the two bars were of opposite contrast. Similarly, Krekelberg and Albright (2000) found that macaques—just like humans—perceived a reversed direction of motion when a stimulus reversed contrast with every displacement in reverse phi. This reversal of perceived direction had a clear correlate in the neural responses of MT cells, which were predictive of the monkey's behavioral decisions.

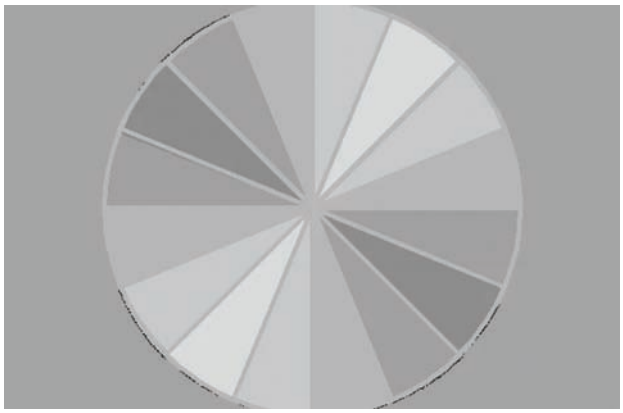
Reverse phi recently found an unexpected real-world application that may benefit some paralyzed people. Spillmann, Anstis, Kurtenbach, and Howard (1997) noted that a random-dot field undergoing counterphase flicker paradoxically appears to move in the same direction as head and eye movements (i.e., opposite to the optic-flow field) due to reverse phi motion caused by apparent pixel movement between successive retinal images. The reversed motion provides a positive feedback control of the display, whereas under normal conditions retinal signals provide a negative feedback. This altered polarity can invoke self-sustaining eye movements akin to involuntary optokinetic nystagmus. Lorenceau (2012) used this display because, although static, it can sustain smooth eye movements in arbitrary directions. After brief training, participants gained

volitional control over smooth pursuit eye movements and could generate digits, letters, words, or drawings at will. For persons deprived of limb movement, this offers a fast, creative, and personal means of linguistic and emotional expression.

BICYCLE SPOKES ILLUSION

Video IV.63-3 shows a sectored disk that rotates clockwise. The thin spokes between the sectors appear to rotate counterclockwise, but in fact the spokes themselves never change their brightness or position. If we gaze at the center of the disk for 20 s and then stop the motion, we see a strong motion after effect, whose direction is clockwise, appropriate to the apparent movement of the spokes, not to the physical jumps of the sectored disk. The spokes must be thin, must lie along the edges of the sectors, and their brightness must be similar to the sectors that they abut. In Video IV.63-3 the spokes have different brightness levels, matched to the sectors they abut, and the disk rocks back and forth through one sector width. Result: All the spokes move in unison, opposite to the direction in which the sectors move.

If one examines the grey levels of the spokes and sectors, one finds that there is one point per rotation when any given spoke first merges with the preceding sector on one side, then with the succeeding sector on the other side. This gives a motion signal that jumps across the tiny width of each spoke. In addition, the movements of the separate spokes are not simultaneous but step clockwise in time around the disk. Gestalt factors group all these sequential movements together, and the entire spoked wheel is perceived as rotating continuously. The tiny counterclockwise spoke movements are more perceptually compelling than the large clockwise sector movements and dominate in the motion aftereffect, showing that they stimulate neural motion detectors more effectively. Possibly, motion sensors with small receptive fields and tuned to small motions are more numerous or more sensitive than sensors tuned to larger motions.



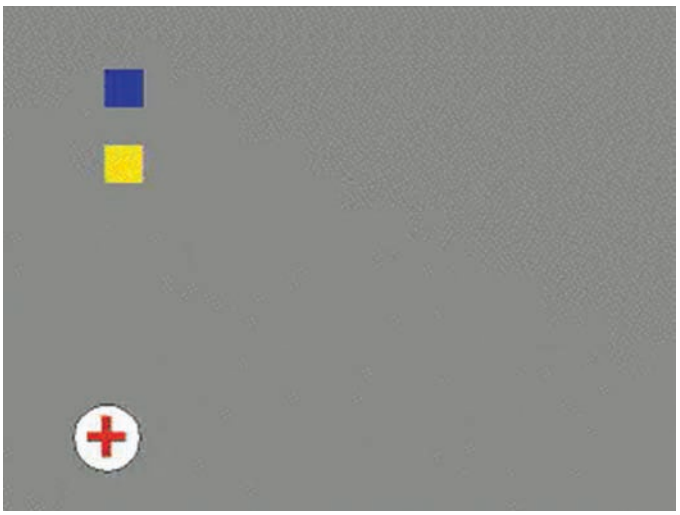
Video IV.63-3. The thin grey spokes never change their luminance or position, but when the sectors jump round clockwise in apparent motion, the spokes appear to rotate counterclockwise. If the motion is then stopped, a clockwise aftereffect of motion is seen—appropriate to the movement of the spokes, not of the sectors. (Video by Stuart Anstis.)

FOOTSTEPS

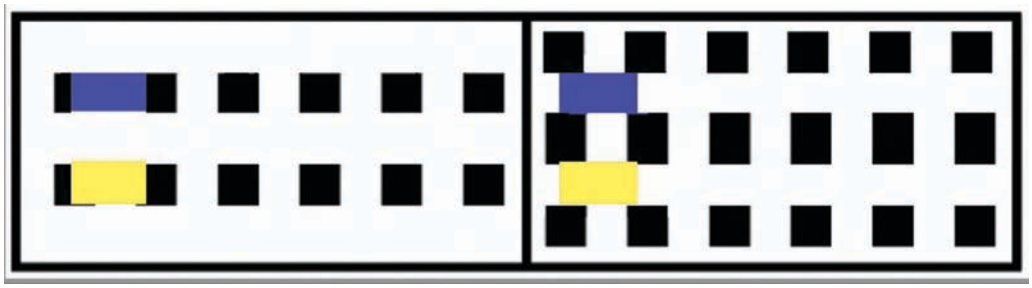
In the footsteps illusion, a dark blue square and a light yellow square, one above the other, move smoothly together across a pattern of vertical black and white stripes. Although the squares actually move at constant speed, their perceived speed varies dramatically. The blue square alternately appears to speed up and slow down, apparently hesitating or even stopping as it traverses each spatial cycle of the grating. Conversely, the yellow square appears to slow down and speed up in counterphase to the blue square (Anstis, 2001, 2003a, 2003b).

As illustrated in Video IV.63-4, the light yellow square appears to slow down on the white stripes, where its edges have low contrast, and speed up on the black stripes, where its edges have high contrast. Conversely, the dark blue square appears to speed up on the white stripes and slow down on the black stripes. Their apparent speeds vary in counterphase, so they look like the two feet of a walking man, one speeding up as the other slows down. Thus the apparent speed of a moving edge depends on its instantaneous contrast against the background. The color of the squares is actually irrelevant, since color has little or no input into motion (Cavanagh, Tyler, & Favreau, 1984)—the blue and yellow simply keep the squares visible even at low contrasts. The footsteps illusion is quite strong if one tracks the moving squares, but it is even stronger if one fixates the stationary cross in Video IV.63-4 so that the squares fall on the peripheral retina.

One can show that the leading and trailing edges are more important than the lateral (top and bottom) edges of the moving bars by selectively removing parts of the background stripes. In a "railroad track" condition, the bars run along a striped "track" of the same vertical height as the stripes (Video IV.63-5, left) so that the leading and trailing edges move over the surround stripes but the lateral edges at top and bottom of the bars do not. This gives a



Video IV.63-4. The two squares move together at constant speed. But fixate the red cross and the two squares seem to speed up and slow down in alternation, like the two feet of a walker. They appear to move fast when they have high contrast and slowly when they have low contrast. (Video by Stuart Anstis.)

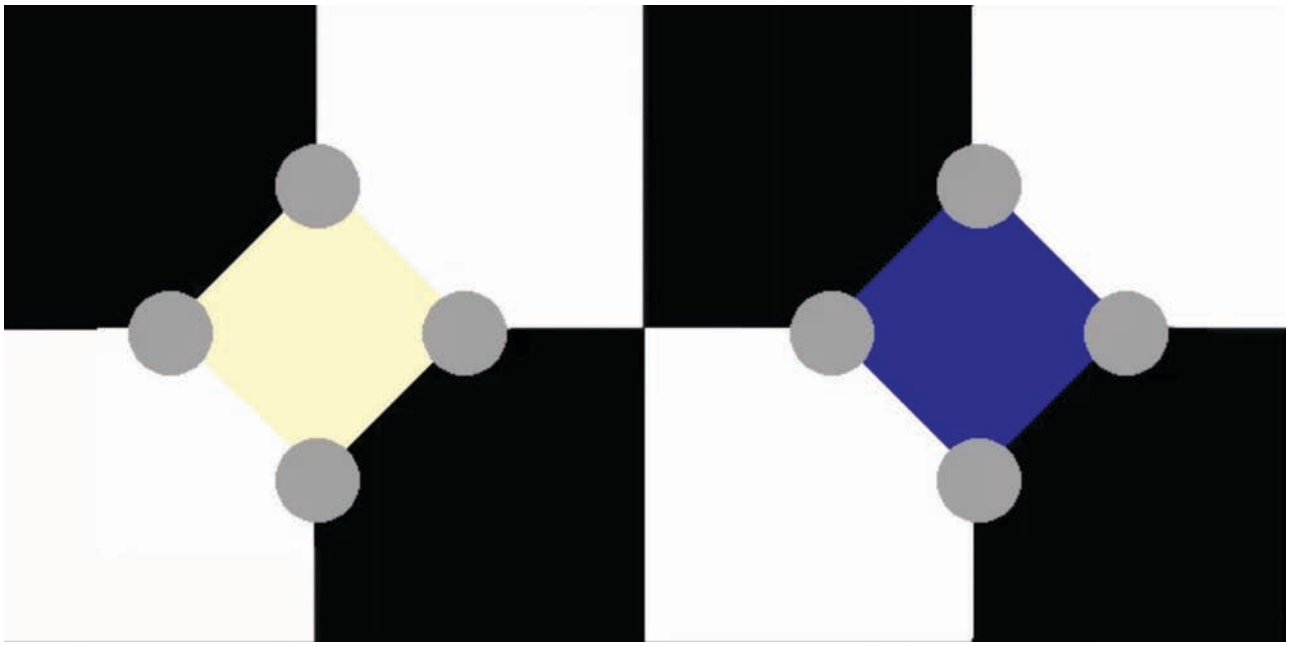


Video IV.63-5. The two bars on the left that run along a "railroad track" show the footsteps illusion, while the two bars on the right that run along a "forest clearing" do not. (Video by Stuart Anstis.)

strong illusion. In a condition that resembles a "clearing in a forest," the squares run along a clear white "track" cut through the surround grating so that the stationary stripes abut only the lateral edges, not the leading and trailing edges, of the squares (Video IV.63-5, right). Now the illusion virtually disappears. I conclude that it is the motion contrast of the leading and trailing edges, not the lateral edges, of the moving bars that produces the footsteps illusion (but see later discussion).

In Video IV.63-6, two footsteps illusions can be placed at right angles and both applied in combination to a single square (Anstis, 2004). Two diamonds, one yellow and one blue, jump up and down through one-tenth of their own diameter, and by varying the contrast along all four of its edges, one can steer its apparent direction (not its speed). The diamond is positioned on a surround of black and white quadrants in such a way that its top-left and bottom-right edges lie on black quadrants. For the light yellow diamond on the left, these two edges have high contrast, so their motion component is subjectively magnified. The other two edges lie on white quadrants. These edges have low contrast, so their motion component is subjectively diminished. The opposite is true for the dark blue diamond on the right. As a result, the motion path appears to be tilted away from the vertical, counterclockwise for the yellow diamond and clockwise for the blue diamond. Measurements show that the perceived direction of motion varies with the 0.6 power of the contrast ratio. This implies that if one edge of the diamond had twice the contrast of the other, the observer could cancel it out by choosing an oblique path that made the lower-contrast edge move through 1.52 times the distance of the higher-contrast edge ($2^{0.6} = 1.52$).

The simplest explanation of the footsteps illusion asserts that perceived speed varies with contrast (Thompson, 1982). Blakemore and Snowden (1999) reviewed the effects of contrast on apparent speed and found that contrast did affect perceived speed for a very wide range of moving stimuli. The footsteps illusion shows that the contrast modulation of speed can be rapid. The bars appear to vary in speed at the temporal frequency with which the bars traverse the stripes, namely 1.8 Hz. Thus the effects of contrast on speed are local in both space and time.



Video IV.63-6. Both diamonds move vertically, but look midway between them and their motion axes appear to be inclined outward. The higher-contrast edges dominate over the lower-contrast edges. (Video by Stuart Anstis.)

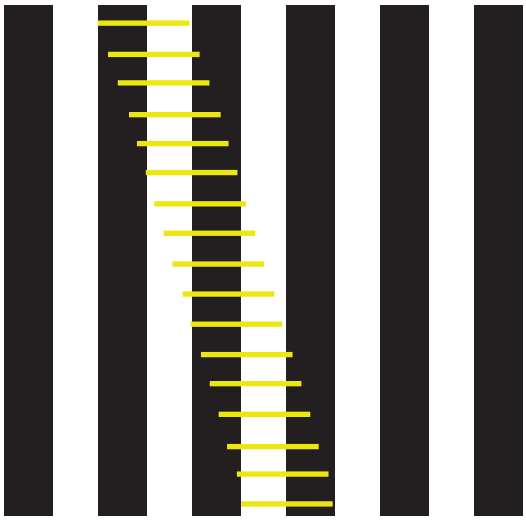


Figure IV.63-1. The Wenceslas illusion—a static form of the footsteps illusion. The straight oblique row of yellow dashes looks curvy. (Figure by Stuart Anstis.)

Each moving bar has the same width as one period of the background grating, so that its leading and trailing edges always lie on the same luminance. Thus when the front and back edges of the dark blue square lie on black stripes, these edges have lower contrast and are harder to see, so the square appears to slow down. When the blue square lies on white stripes, its edges have higher contrast and are easier to see, so the square appears to speed up. The opposite is true for the light yellow square.

However, Howe, Thompson, Anstis, Sagreiya, and Livingstone (2006) made some further observations that suggest that this explanation is too simple and that the contrast of the lateral (top and bottom) edges also plays a part. The top and bottom edges cannot add any motion information—all they can do is dilute the motion. Howe et al. suggest that the signal from each edge is weighted by its contrast, so the total perceived motion is the weighted average of the signals from each of the four edges. Thus the footsteps illusion is caused by the variations in luminance contrast at the leading and trailing edges of each bar, relative to the variations in luminance contrast at the lateral edges of the bar.

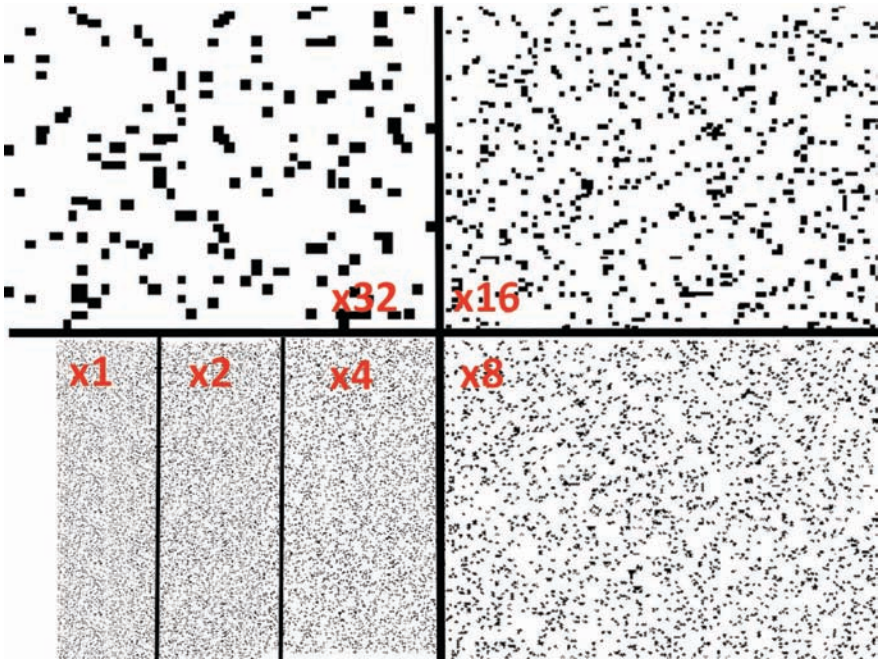
What could be the neural substrate of this dependence of perceived speed on contrast? It seems intuitively likely that motion-sensitive neurons in the brain would signal less vigorously when stimulated by moving targets whose contrast is low and respond more briskly to higher-contrast moving patterns, and Thiele, Dobkins, and Albright (2000) have recently discovered just such responses in single neurons in macaque visual area MT (see their Fig. 4). These neural findings could go a long way in explaining the footsteps illusion.

Similar illusions exist for position as well for motion. Static bars in Figure IV.63-1 show an analogous illusion of position, which is known as the Wenceslas illusion (Thompson & Anstis, 2005: see also Sunaga, Sato, Arikado, & Jomoto, 2008).

The ends of the bars lie along obliques that are actually straight but appear to undulate. This is consistent with the fact that the same visual processing stream that handles motion perception is also responsible for position perception (Livingstone & Hubel, 1987).

ZIGZAG MOTION

If a sparse pattern of randomly scattered dots makes a sequence of small jumps to the right, one naturally sees motion to the right. If it makes a sequence of jumps downward, but the jumps are "too big" for the visual system to take in, one sees not motion downward but dots jumping



Video IV.63-7. All the moving random-dot stimuli are identical, except that they are shown at magnifications of x1, 2, 4, 8, 16, 32. Yet the smaller magnifications seem to move downward, while the identical stimuli at higher magnifications appear to move to the right. The reason is each random-dot field makes small horizontal jumps, alternating with large downward jumps, as if descending a steep staircase. At small magnifications, these motions are seen veridically. But at large magnifications, the large jumps give only weak signals of motion, and the visual system loses track of which dot is which. The result is only the small rightward jumps are visible at the highest magnifications. (Video by Stuart Anstis.)



Video IV.63-8. A striped disk moves vertically up and down across oblique stripes. In foveal vision this is seen veridically. But at increasing retinal eccentricities, the disk's path looks increasingly oblique, until in far peripheral vision the disk appears to move parallel to the stripes. (Video by Stuart Anstis.)

about randomly. There is a limit to the jump size that the visual system can tolerate.

Now look at the set of movies embodied in Video IV.63-7. Each video contains the identical set of moving random dots, which differ only in their relative magnifications. These are accurately labeled as x1, x2, x4, x8, x16, and x32. When they are set in motion, most people see the dots in the bottom row as mostly moving downward and those in the top row as moving to the right, with perhaps doubtful directions of motion for the x8 and x16 dots. In fact, all the dots are moving in exactly the same way (except magnified accordingly to the dot sizes). To verify this, inspect Video IV.63-7 from different viewing distances, especially doubtful cases such as x8. Viewed from close up, which increases the magnification, the x8 dots appear to move to the right. Viewed from further away, which decreases their magnification, the very same dots appear to move downward.

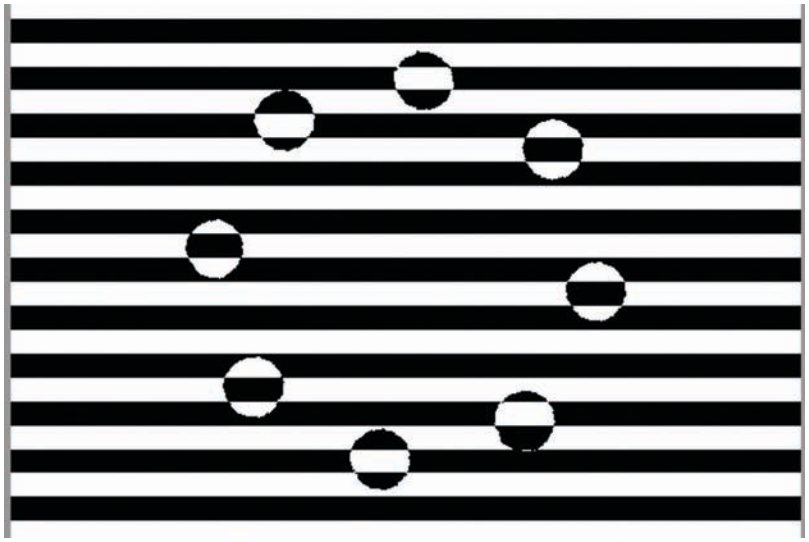
How is this possible? Each set of random dots moves rigidly, making a small jump to the right followed by a downward jump 10 times larger, in an endless alternation. So the dots are descending a very steep staircase. At small magnifications such as x1 or x2, one sees the alternating jumps, but the sideways jumps are almost too small to notice, and the overall path of the dots is almost vertically downward, toward a direction between 5 o'clock and 6 o'clock. However, for the large magnifications such as x32 or x16, the large jumps are too large for the visual system to respond to adequately, and observers notice only the rightward jumps. The large jumps exceed Dmax (Baker & Braddick 1985), correspondence is lost, and the observer loses track of which dot is which.

We conclude that, for apparent motion, less is more. Small jumps stimulate large numbers of motion-selective neurons that are tuned to small jumps and slow motion, whereas there are very few motion-selective neurons tuned to very large jumps.

THE FURROW ILLUSION: PERIPHERAL MOTION

The retinal periphery has far less cortex at its disposal than the fovea, so it is forced to use highly economical coding. This involves poor acuity plus, it seems, an inability to combine local neural signals of motion into the percept of complete moving objects.

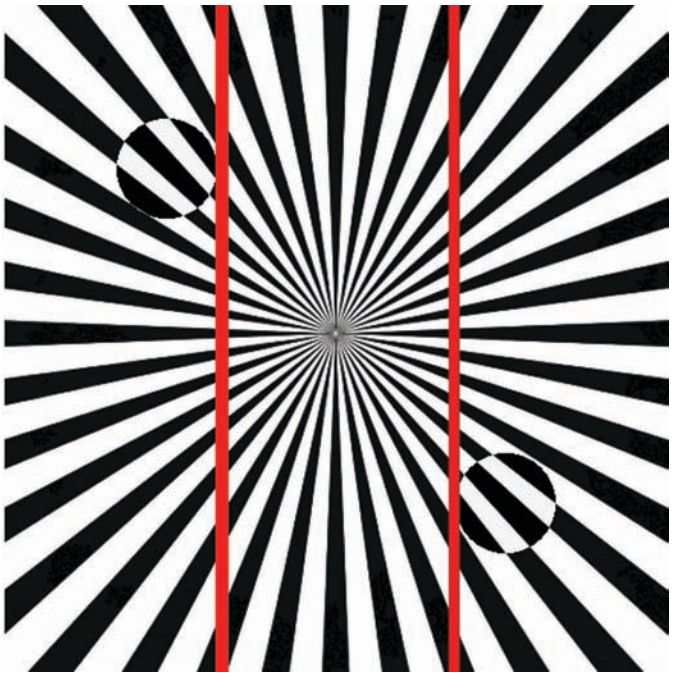
Video IV.63-8 shows a spot moving vertically across an oblique grating. The spot is a "negative lens" because this gives more salient intersections, which seem to bolster the illusion. In foveal vision one sees the vertical motion veridically, but in increasingly peripheral vision the motion path seems to diverge more and more from the vertical until at large eccentricities it appears to move along a 45° trajectory, parallel to the background grating.



Video IV.63-9. A circle of eight disks rotates clockwise against a background of stripes that are alternately vertical and horizontal. In foveal vision this is seen veridically. But in peripheral vision the disks appear to slip and slide around an ellipse that is roughly parallel to the background stripes. (Video by Stuart Anstis.)



Video IV.63-10. The moving horizontal bar is always the same size. But it appears to change dramatically in its apparent length and depth, even in foveal vision but particularly in peripheral vision. (Video by Stuart Anstis.)



Video IV.63-11. In the Hering (1861) illusion the red vertical lines are parallel but appear to bow slightly outward like the sides of a barrel. But the Furrow illusion is *not* simply a dynamic version of this. When the striped disks move up and down, always kissing the red lines, they appear to bow strongly inward (pincushion distortion), opposite to the direction of the Hering illusion. The Hering illusion shows orientation contrast while the Furrow illusion shows orientation assimilation. (Video by Stuart Anstis.)

In Video IV.63-9, spots move around in a circle against a background of stripes that periodically change their orientation. In foveal vision the moving spots are seen veridically, but in the periphery they appear to slide around an elliptical path that aligns with the background stripes as these alternate between horizontal and vertical.

In Video IV.63-10, the moving horizontal bar never changes its length or width, but it appears to change its size and depth dramatically, especially in peripheral vision. In Video IV.63-11, two spots move up and down along vertical paths, kissing two static vertical red lines. Note that these lines appear to bow slightly outward in barrel distortion, owing to the Hering (1861) illusion. But in peripheral vision the paths of the spots appear to bow strongly inward, opposite to the Hering distortion.

Thus while the Hering illusion is an instance of orientation contrast, this "furrow illusion" (Anstis, 2013) is an instance of orientation assimilation. The furrow illusion is consistent with the idea that the peripheral visual field responds adequately to the moving intersections shown in Video IV.63-8 but is unable to integrate these local intersection signals into the percept of a complete moving object (Braddick, 1993). I conclude that the retinal periphery sacrifices this top-down ability in the interests of economy.

REFERENCES

Adelson, E. H., & Bergen, J. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America, 2*(2), 284–299.

Anstis, S. M. (1970). Phi movement as a subtraction process. *Vision Research*, *10*, 1411–1430.

Anstis S. (2001). Footsteps and inchworms: Illusions show that contrast affects apparent speed. *Perception*, *30*, 785–794.

Anstis, S. (2003a). Levels of motion perception. In L. Harris & M. Jenkin (Eds.), *Levels of perception* (pp. 75–99). New York: Springer.

Anstis S. (2003b). Moving objects appear to slow down at low contrasts. *Neural Networks*, *16*, 933–938.

Anstis S. (2004). Factors affecting footsteps: Contrast can change the apparent speed, amplitude and direction of motion. *Vision Research*, *44*, 2171–2178.

Anstis, S. (2013). Contour adaptation. *Journal of Vision, 13*(2). http://jov.arvojournals.org/article.aspx?articleid=2193818

Anstis, S., & Mather, G. (1985). Effects of luminance and contrast on direction of ambiguous apparent motion. *Perception*, *14*(2), 167–179.

Anstis, S. M., & Rogers, B. J. (1975). Illusory reversal of visual depth and movement during changes of contrast. *Vision Research*, *15*, 957–961.

Anstis, S., Smith, D. R., & Mather, G. (2000). Luminance processing in apparent motion, Vernier offset and stereoscopic depth. *Vision Research*, *40*(6), 657–675.

Baker, C. L. Jr., & Braddick, O. J. (1985). Eccentricity-dependent scaling of the limits for short-range apparent motion perception. *Vision Research, 25*(6), 803–812.

Blakemore, M. R., & Snowden, R. J. (1999). The effect of contrast upon perceived speed: A general phenomenon? *Perception*, *28*, 33–48.

Blakeslee, B., Pasieka, W., & McCourt, M. E. (2005). Oriented multiscale spatial filtering and contrast normalization: A parsimonious model of brightness induction in a continuum of stimuli including White, Howe and simultaneous brightness contrast. *Vision Research, 45*(5), 607–615.

Bours, R. J. E., Kroes, M. C. W., & Lankheet, M. J. M. (2007). The parallel between reverse phi and motion aftereffects. *Journal of Vision*, *7*(11), 8.

Bours, R. J., Kroes, M. C., & Lankheet, M. J. (2009). Sensitivity for reverse phi motion. *Vision Research*, *49*(1), 1–9.

Braddick, O. (1993). Segmentation versus integration in visual motion processing. *Trends in Neurosciences*, *16*(7), 263–268.

Cavanagh, P., Tyler, C. W., & Favreau, O. E. (1984). Perceived velocity of moving chromatic gratings. *Journal of the Optical Society of America*, *1*, 893–899. http://jov.arvojournals.org/article.aspx?articleid=2193004

Hering, E. (1861). *Beiträge zur physiologie: I. Zur Lehre vom Ortssinne der Netzhaut*. Leipzig, Germany: Engelmann.

Howe, P. D. L., Thompson, P. G., Anstis, S. M., Sagreiya, H., & Livingstone, M. (2006). Explaining the footsteps, belly dancer, Wenceslas and kickback illusions. *Journal of Vision, 6*(12), 1396–1405.

Krekelberg, B & Albright, T.D., (2005). Motion mechanisms in macaque MT. *J Neurophysiol*. 93(5), 2908–2921.

Krekelberg, B & Lappe, M. (2000). A model of the perceived relative position of moving objects based upon a slow averaging process. *Vision Research , 40*, 201-215.

Livingstone, M., & Conway, N. (2002). Responses of V1 neurons to reverse phi stimuli. *Journal of Vision*, *2*(7), 127.

Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience, 7*, 3416–3468.

Lorenceau, J. (2012). Cursive writing with smooth pursuit eye movements. *Current Biology*, *22*(16), 1506–1509.

Reichardt, W. (1986). Processing of optical information by the visual system of the fly. *Vision Research, 26*, 113–126.

Spillmann, L., Anstis, S., Kurtenbach, A., & Howard, I. (1997). Reversed visual motion and self-sustaining eye oscillations. *Perception*, *26*(7), 823–830.

Sunaga, S., Sato, M., Arikado, N., & Jomoto, H. (2008). A static geometrical illusion contributes largely to the footsteps illusion. *Perception, 37*(6), 902–914.

Thiele, A., Dobkins, K. R., & Albright, T. D., (2000). Neural correlates of contrast detection at threshold. *Neuron, 26*, 715–724.

Thompson, P. (1982). Perceived rate of movement depends on contrast. *Vision Research*, *22*, 377–380.

Thompson, P., & Anstis, S. (2005). Retracing our footsteps: A revised theory of the footsteps illusion. *Journal of Vision*, *5*(8), 929.

Walls, G. (1942). *The vertebrate eye and its adaptive radiation*. Bloomfield Hills, MI: Cranbrook Institute of Science.

Watson, A. B., & Ahumada, A. J. Jr. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A*, *2*, 322–341.

Wehrhahn, C. (2006). Reversed phi revisited. *Journal of Vision*, *6*(10), 2.

White, M. (1981). The effect of the nature of the surround on the perceived lightness of grey bars within square-wave test gratings. *Perception, 10*(2), 215–230.

Chapter 64
High-Level Organization of Motion

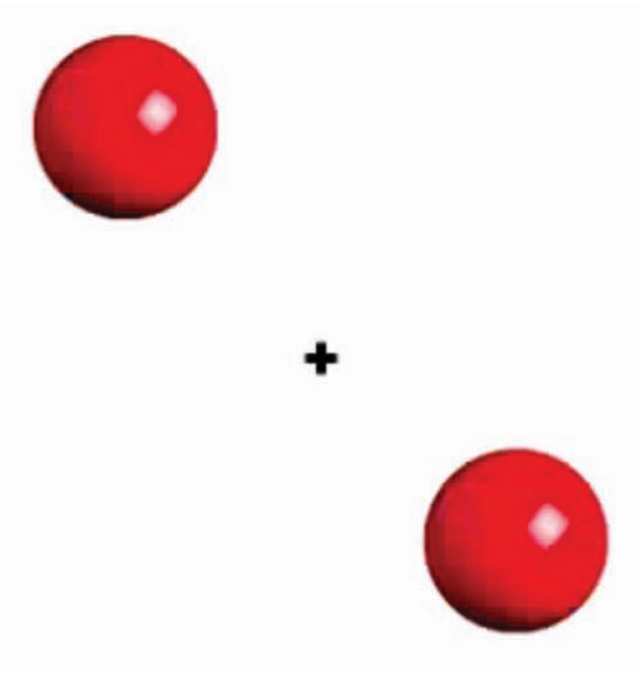
Ambiguous, Primed, Sliding, and Flashed

Stuart Anstis

AMBIGUOUS APPARENT MOTION

In stroboscopic or apparent motion, a spot that jumps back and forth between two positions appears to be moving continuously back and forth, probably because it is adequate, although not optimal, to stimulate neural motion detectors.

Ambiguous apparent motion stimuli can be constructed, as in the two-frame Video IV.64-1. Frame 1 flashes up two red spots at opposite corners of an imaginary square. Frame 2 flashes up spots at the other two corners (Gengerelli, 1948). What do you see? You may see horizontal apparent motion, back and forth along the top and bottom of the square, or you may see vertical apparent motion, back and forth up and down the left and right sides of the square. The stimulus is ambiguous and supports both of these interpretations equally, and your percept is likely to change back and forth pseudo-randomly between horizontal and vertical motion, even while the stimulus stays the same. If you attempt to influence which you see by means of "willpower," at first it seems possible, but then the percept seems to escape and show, as if it had a will of its own. Hiding two spots removes the ambiguity. Fixate the central cross and cover the two bottom spots with your hand. You will now see unambiguous horizontal motion along the top of the square. Remove your hand and the horizontal motion continues to dominate—for a while. Similarly, covering up the two right-hand spots forces vertical motion up and down the left side of the square, which continues when all four spots are uncovered—for a while. This display is bistable, just as a Necker cube is, and the changing percepts from an unchanging stimulus are a way to see your brain's computations at work.

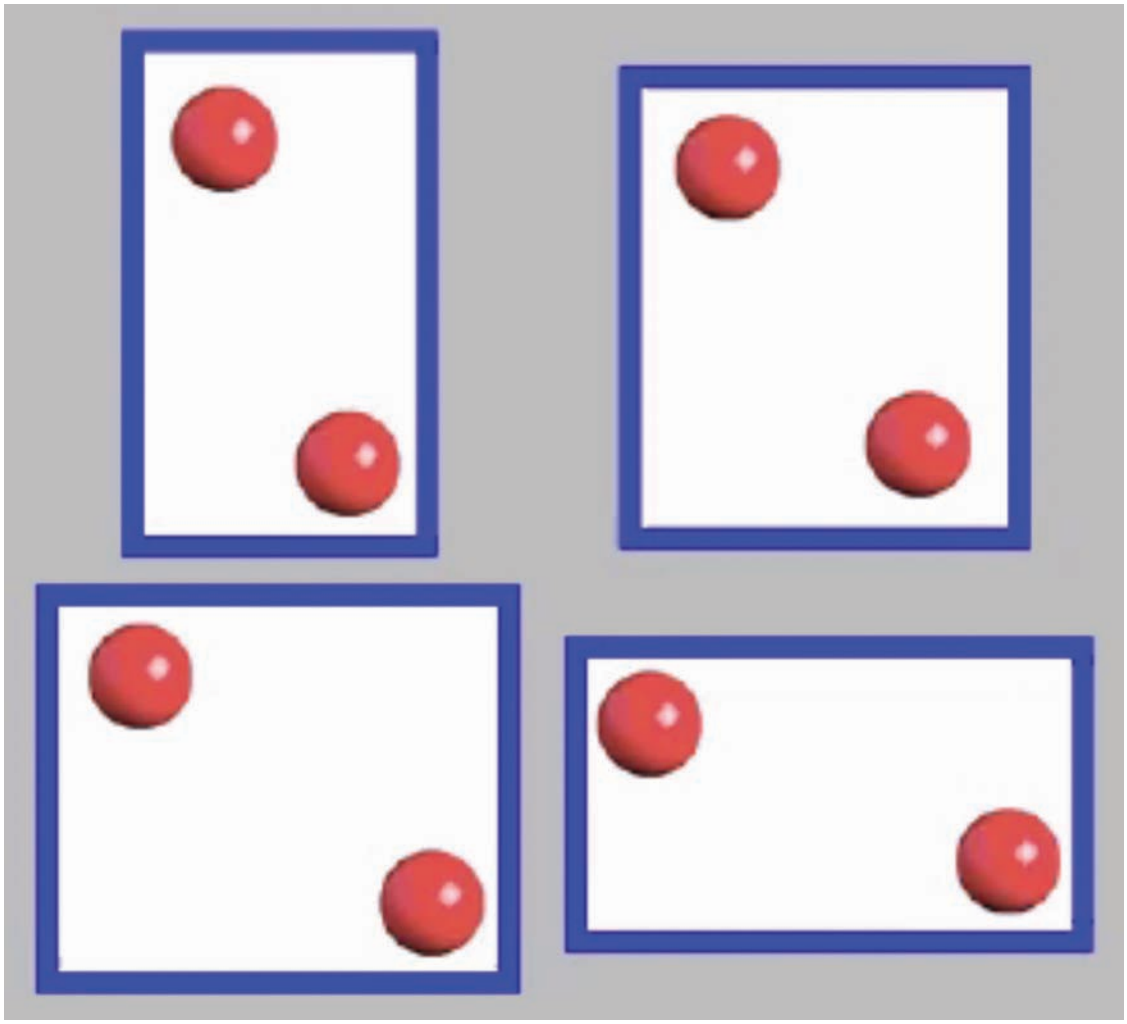


Video IV.64-1. Ambiguous apparent motion. One can see two spots jumping back and forth, either horizontally or vertically but not both at once. The ambiguous motion flips back and forth randomly between the two directions over time. Hiding the top two spots with one's hand primes horizontal motion of the two bottom spots, which generally persists for a while after the hand is removed. This stimulus is like a "dynamic Necker cube." (Video by Stuart Anstis.)

Proximity can reduce ambiguity. In Video IV.64-2 one usually sees motion along the short edges of the rectangles: horizontal motion along the short top and bottom edges of the tall rectangles and vertical motion along the short vertical sides of the wide rectangles. Thus the visual system prefers to minimize path lengths.

There is a permanent perceptual bias in favor of seeing up-and-down motion between vertically aligned spots, which are both contained in one cortical hemisphere, as opposed to side-to-side motion of horizontally aligned spots on either side of the midline, which requires collaboration between the two hemispheres. For this reason it is better to put the spots at the corners of an imaginary diamond instead of a square.

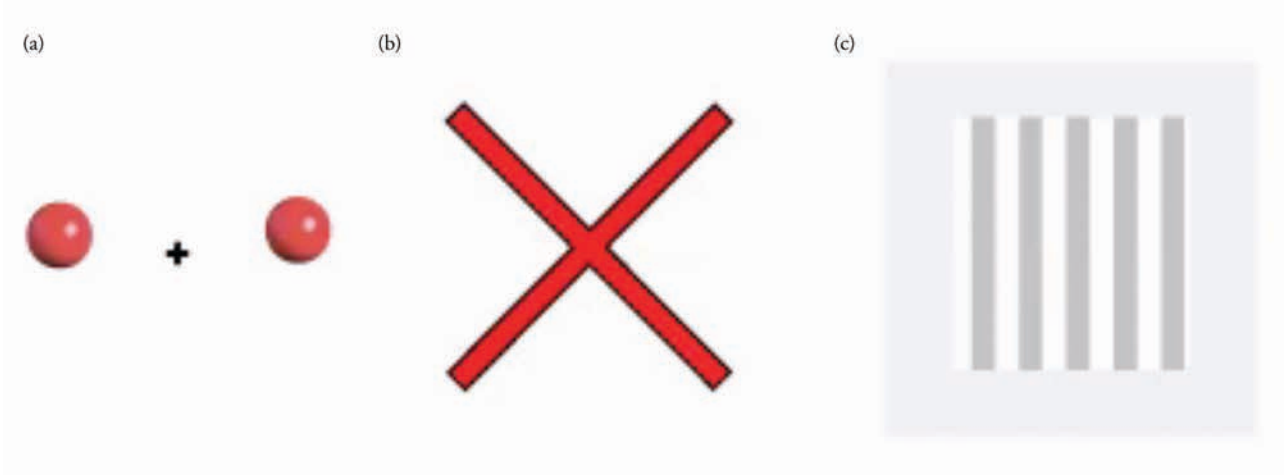
Long-range spatial interactions affect the perceived apparent motion (Ramachandran & Anstis, 1983). If several such bistable figures are randomly scattered on a screen and presented simultaneously, then observers



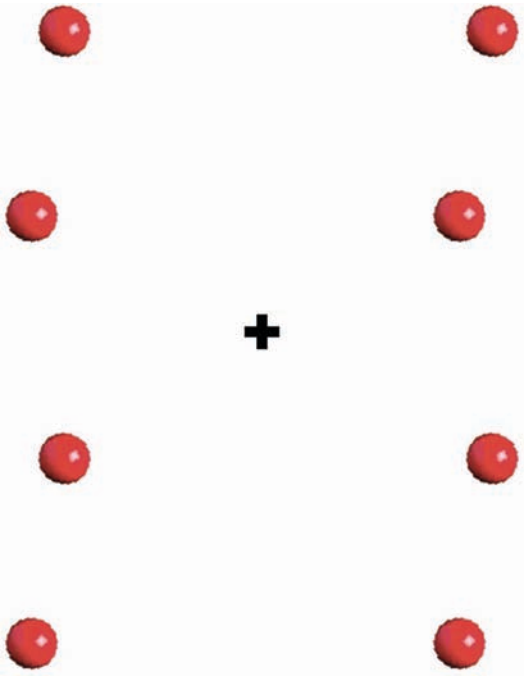
Video IV.64-2. Ambiguous apparent motion is often driven by proximity, so that motion is seen between nearest neighbors. Thus the perceived motion is usually horizontal in the top left panel and vertical in the bottom right panel. (Video by Stuart Anstis.)

always see the same motion axis in all of them (all move vertically or all move horizontally), suggesting the presence of global field-like effects for resolving the motion ambiguity.

Temporal interactions can prime the perceived direction of apparent motion if a prior frame is added containing two more spots in line with two parallel sides of the diamond, as in Video IV.64-3a (Anstis & Ramachandran, 1987). The



Video IV.64-3. Ambiguous apparent motion can be primed to go in the direction of a preliminary nonambiguous motion. (Video by Stuart Anstis.)



Video IV.64-4. Local versus global motion. In this ambiguous stimulus, four pairs of spots rotate locally about their common centers. Following an inspection period, the percept can change suddenly into two large squares of spots, each following a global circular path. Global tends to predominate more over local motion as time passes. (Video by Stuart Anstis.)

upper spot now always moves along those sides, down to the right, not the left, because the visual system prefers to see motion in a straight line rather than along an L-shaped trajectory. Visual inertia is still seen when the priming dots are presented to one eye and the test dots to the other, so the effect must be partly central. This "visual inertia" is also shown in Video IV.64-3b (Anstis & Ramachandran, 1987). In (b), an upright cross that suddenly jumps through 45° is ambiguous, since the apparent motion is equally likely to be clockwise (CW) or counterclockwise (CCW). But adding an initial cross that jumps unambiguously CW first biases the motion to be always CW.

In the same way, Pinkus and Pantle (1997) noted that the perceived motion of a vertical sine-wave luminance grating that undergoes an abrupt 180° phase shift (motion step) is ambiguous. The grating sometimes appears to move rightward, sometimes leftward. However, when the 180° step follows closely upon an unambiguous grating step of 90° to the right, as in Video IV.64-3c, the 180° step also appears to be to the right.

LOCAL/GLOBAL MOTION

Ambiguous static pictures such as the Necker cube or Rubin's face/vase are very well known, but I shall now describe some ambiguously moving patterns, which the visual system can organize perceptually in different ways corresponding to different aspects of the movement (Anstis & Kim, 2011). The Gestalt notion of "common fate" needs to be refined to account for these ambiguities.



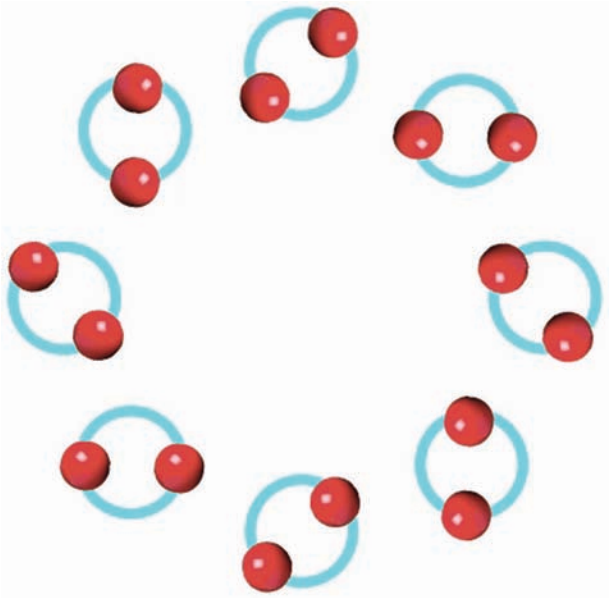
Video IV.64-5. Local versus global motion. The motion of this stimulus stably remains local as two lovers gaze into each other's eyes. (Video by Stuart Anstis.)

Video IV.64-4 shows four pairs of red spots, each pair rotating about their common center. But if one inspects this pattern for 10 to 20 s, the pattern can suddenly reorganize and be seen as two large squares drifting over each other. I call this "local motion" when each spot groups with its single nearest neighbor, forming four pairs, and "global motion" when the spots reorganize into two quartets. During prolonged inspection the local and global motion tend to alternate, with the global motion gradually becoming more frequent. Proximity or similarity to local or distant partners can bias the percept toward local or global motion.

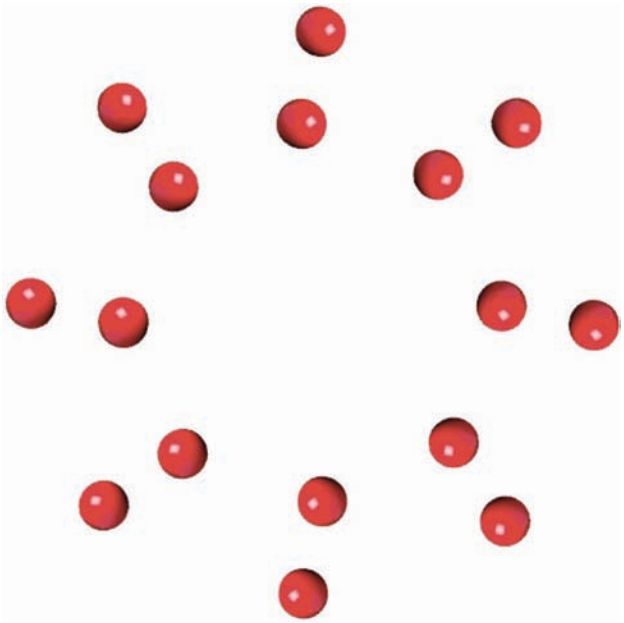
Thus Video IV.64-5 tends to be seen locally as four pairs of inward-looking faces. Setting members of each pair closer together or farther apart promotes local or global motion, respectively, and increasing the number of spots within a group from two to three or four makes the motion look more local (not shown).

In Video IV.64-6 one generally sees eight local pairs of spots moving around circular orbits, especially if one looks straight at a pair. When the circles are removed, as in Video IV.64-7, one generally sees two global, interdigitated octagons, especially if one looks at the middle of the figure.

Local and global motion are two different and incompatible solutions to the problem of binding dots into groups. They are incompatible because it is impossible to see the same dots as partaking in local and global groups simultaneously. I suspect that local motion is preattentive, whereas global motion is attentive. Local and global perceptual organizations are like scientific hypotheses that attempt to group the most data under the fewest and most general laws. Thus global motion would be a more generalized perceptual achievement because it organizes more spots into fewer groups.



Video IV.64-6. Local versus global motion. Gaze at any one of these rotating pairs and the motion tends to remain local, aided by the blue circles that show the local trajectories. (Video by Stuart Anstis.)



Video IV.64-7. Local versus global motion. Fixate the center of this video, which resembles Video IV.64-6 without the circles. One then sees two intertwined global octagonal arrays. (Video by Stuart Anstis.)

FLYING BUGS

The flying bugs illusion is an example of induced movement or relative movement (Duncker, 1929/1955). A familiar example is the moon, which appears to be sailing to the right when clouds drift across it moving to the left. The flying bugs illusion convincingly demonstrates that a moving background can strongly influence the perceived paths of small moving objects.

In Video IV.64-8, two bugs fly around CW circular orbits, in counterphase so that when one bug is at 12 o'clock the other is at 6 o'clock. When a large background is added, also moving along a CW path without rotating and in step with the left-hand bug, there is a strong size illusion that the left-hand bug is flying around in much smaller circles than the right-hand bug, as shown in Video IV.64-9.



Video IV.64-8. Flying bugs: Control condition. Two bugs fly around in circles. Note: They follow exactly these same orbits in Videos IV.64-9 and IV.64-10. (Video by Stuart Anstis.)



Video IV.64-9. Flying bugs: Size illusion. The left-hand bug seems to fly in much smaller circles, because it moves in step with the background so it has no motion relative to the background. The right-hand bug has the same physical motion as the left-hand bug, but it moves strongly relative to the background, so its perceived motion is increased. (Video by Stuart Anstis.)

This is because the left-hand bug moves exactly with the background, and therefore has zero movement *relative* to the background. The right-hand bug is out of phase with the background so it describes circles about twice as large as before *relative* to the background, although the orbits of both bugs have the same physical size.

If the background reverses direction and moves CCW, a shape illusion is observed (Video IV.64-10). The left-hand bug seems to move mostly from side to side, while the right-hand bug appears to move mostly up and down. This is because the circular motion of the bugs has a vertical (sin) component and a horizontal (cos) component. For the left-hand bug, the vertical component is in phase with the background so is mostly ignored, while its horizontal component moves in opposition to the background so is perceptually enhanced. The opposite is true for the right-hand bug.

The background motion need not be continuous. In Video IV.64-11 it consists of a series of short-lived unrelated textures, with the first texture moving briefly toward 1 o'clock that is then abruptly replaced by a different texture moving toward 2 o'clock, then a third texture moving toward 3 o'clock, and so on. The moving textures look, and are, jerky and disconnected, but in combination they impart a smooth induced movement into the bugs' orbit, so this induced movement is local with respect to space.

Zivotofsky (2004) moved a single red test dot horizontally left or right while a dense background set of black dots on a white background moved vertically up or down. When the background-inducing dots moved up (down), the truly horizontally translating test dot appeared to drift at an angle down (up) from the horizontal. Farrell-Whelan, Wenderoth, and Wiese (2012) measured the angular function of Duncker's effect by moving the red spot in one direction and the random-dot background in a series of other directions. They found that induced movement was strongest when the two directions of motion differed by 30°, similar to the inducing angle that has been found to maximize other direction illusions.

Rhesus monkeys behave as if they perceive induced movement (Zivotofsky, Goldberg, & Powell, 2005). All of these results can be modeled by a simple algorithm in which the average motion of all the points in the whole field put together is subtracted from each individual point (Burt & Sperling, 1981).

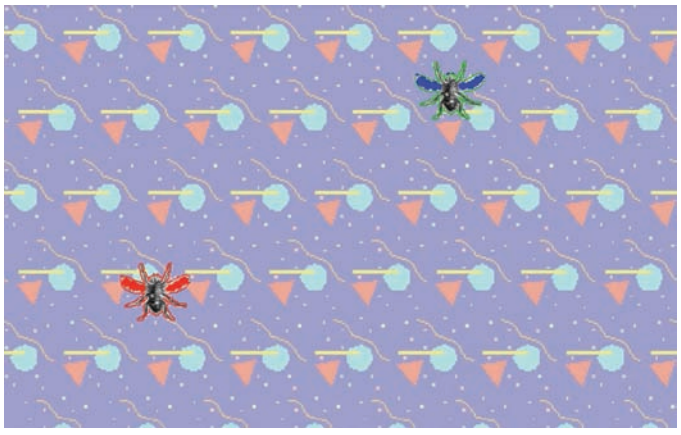
CHOPSTICKS ILLUSION

The closing blades of a pair of scissors form an intersection that can cut one's fingernails, and whose position can be perceived accurately, although it is not really an object. But moving intersections are sometimes grossly misperceived. In the chopstick illusion a vertical and a horizontal line overlap to form a cross. Each line follows a CW circular path without rotating, like the sponge in the hand of a window cleaner. The vertical and horizontal lines move in counterphase, with one line being at 6 o'clock when the other line is at 12 o'clock (Video IV.64-12a).



Video IV.64-10. Flying bugs: Shape illusion. The left-hand bug seems to fly from side to side, because the vertical component of its circular motion matches the vertical motion of the background so is perceptually cancelled out. Conversely, the horizontal component of its circular motion is strongly against that of the background so is perceptually enhanced. The opposite is true for the right-hand bug, which seems to fly up and down. (Video by Stuart Anstis.)

The illusion arises in the central intersection where the lines cross. This intersection actually moves CCW around a circle, but it is incorrectly perceived as moving CW. The argument goes in two stages. First, the middle of each line gives no unambiguous information about its motion—that information comes from the tips or terminators of the lines. Second, the CW motion is in fact the relative movement of each of the two lines with respect to the other (Bruno & Bertamini, 2013). A tiny observer standing on either of the lines would see the other line as rotating CW around him or her.

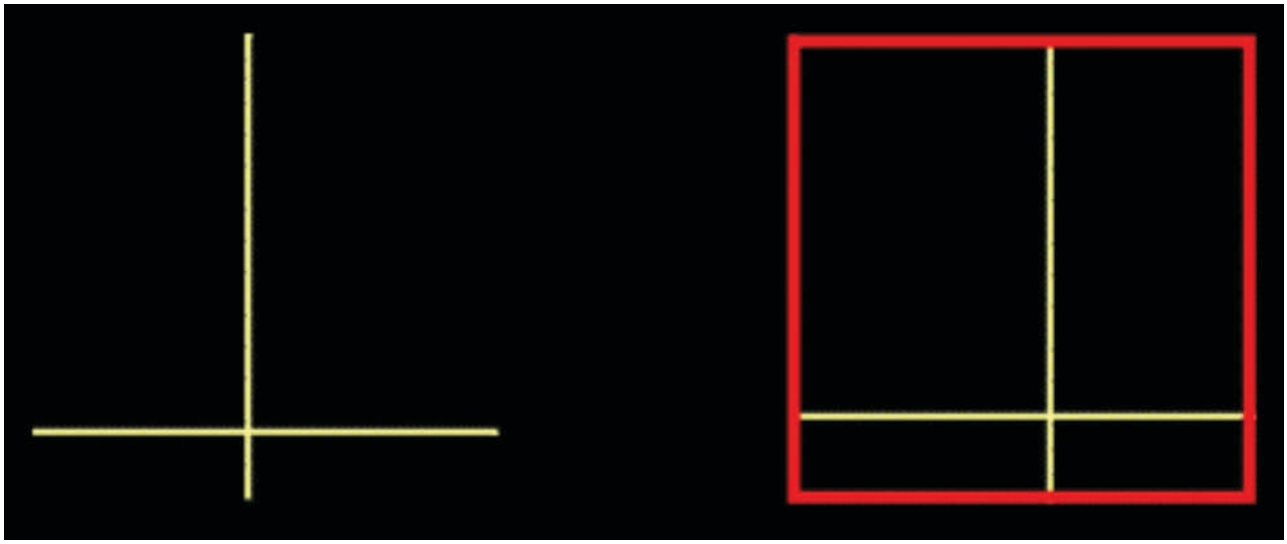


Video IV.64-11. Bursts of background motion still make the right-hand bug's orbit look much bigger than the left-hand bug's. (Video by Stuart Anstis.)

If a floating outline square is added that just touches all four tips of the two rotating lines, the illusion is abolished and the intersection is correctly seen as a rigid X-intersection moving CCW, even though the terminators are fully visible (Video IV.64-12b). The reason is that the terminators are now seen not as informative true ends of lines but as the "extrinsic" visible parts of extended lines that continue behind the square frame and give no clue as to the real motion of the lines. Intrinsic terminators signal the physical end of an edge of an object, whereas extrinsic terminators are caused by occlusion. Intrinsic terminators provide an unambiguous motion signal regarding the true direction of object motion, while extrinsic terminators provide a locally ambiguous signal that must be suppressed for accurate motion computation (Liden & Pack, 1999; Shimojo, Nakayama, & Silverman, 1989).

Thus when the terminators are perceived as intrinsic, they drive the chopstick illusion, but when they are perceived to be extrinsic—a consequence of occlusion by the aperture—they are ignored. Stoner and Albright (1994) comment,

> "[Results like these] are particularly important because they emphasize the ability of cues for feature classification to "act at a distance," governing the integration of motion signals at locations in the image where segmentation cues are either absent or ambiguous. Models of motion signal integration must thus provide for the influence of such non-local information." (pp. 279–280)



Video IV.64-12. The yellow vertical and horizontal rods follow counterphase CW paths. The central intersection of the rods seems to move CW in the left-hand display and CCW in the right-hand display. Actually it moves CCW in both displays. (Video by Stuart Anstis.)

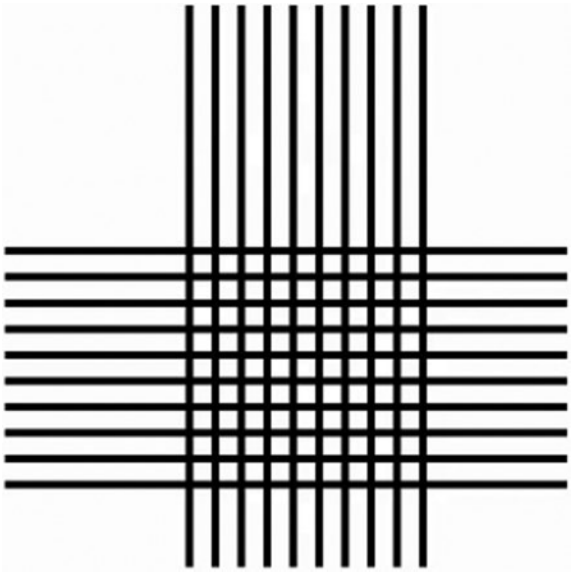
Regarding Numbers of Intersections Versus Tips

In Video IV.64-12 there are four line terminators but only one intersection. We tested whether the tips won out by sheer force of numbers by increasing the number of lines. In Video IV.64-13 there are 10 vertical and 10 horizontal lines, which gives 40 terminators and 100 (= 10^2) intersections. Now the intersections outnumber the terminators by 2.5:1, yet a chopstick illusion is still seen. We conclude that the chopstick illusion is not caused by mere force of numbers.

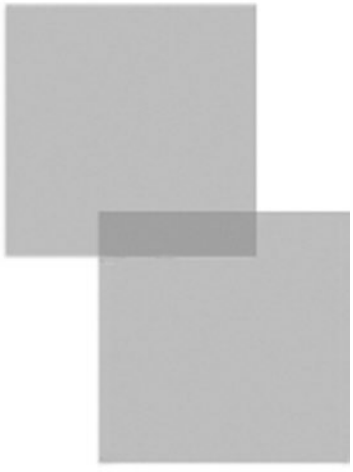
Instead of a horizontal and a vertical line, one can use two overlapping transparent squares, with each square moving CW. The places where the sides of the two squares intersect in Video IV.64-14 actually move CCW but appear to move CW. It is not necessary for the squares to obey the rules of transparency (Metelli, 1974): the illusion still holds whatever the brightness of the region of overlap (not shown).

The shape of the aperture makes a difference. In Video IV.64-15a the moving lines are viewed through an oblique slot in a *visible* black mask. This gives no illusion, and the central intersection looks like a rigid cross circling CCW. But Video IV.64-15b shows an oblique slot in an *invisible* white mask.

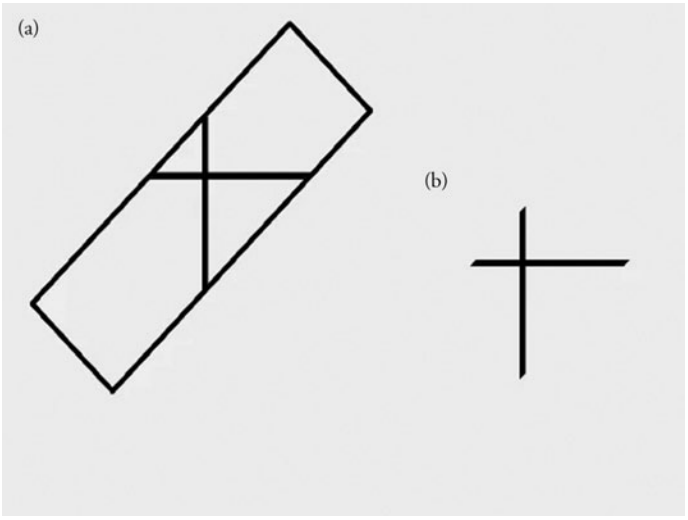
In (a), the mask is visible and the intersection is correctly seen as rotating CCW, since the terminators are treated as extrinsic and ignored. In (b), the oblique mask is invisible and the illusion returns; the central intersection appears to move obliquely, parallel to the oblique paths of the terminators, because these terminators are treated as intrinsic and are taken into account.



Video IV.64-13. Even with many chopsticks the illusion is still seen, even though the 100 intersections far outnumber the 40 line tips (terminators). (Video by Stuart Anstis.)



Video IV.64-14. The chopstick illusion is still seen with two moving transparent squares. The intersections of the squares move CCW but still seem to move CW. (Video by Stuart Anstis.)



Video IV.64-15. A cross, seen through an obliquely oriented window, moves CCW around a circle. When the edges of the window are visible (left), the motion is seen correctly. But when the edges are invisible (right), the lines and the intersections all appear to slide obliquely, parallel to the long sides of the window. (Video by Stuart Anstis.)

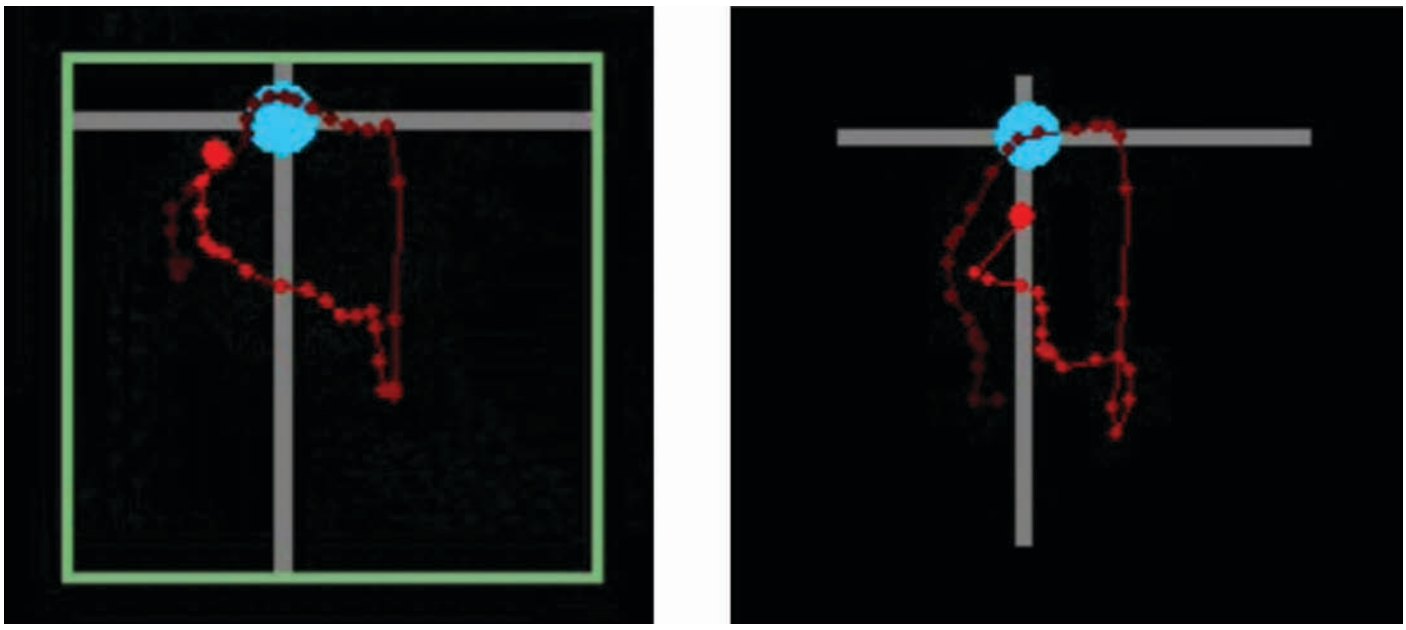
Eye Movements

Observers have great difficulty in tracking the movements of a sliding intersection. It moves CCW but appears to move CW, and eye movements show 10 times the errors of a control condition (not shown), in which observers simply track a cross that circles CW (Anstis & Ito, 2010). Video IV.64-16 shows the same stimuli as Video IV.64-12, with the real-time attempted eye-tracking movements of an observer superimposed in red. The blue disk that marks the intersection in the movie was not shown to the observer. Notice that the observer tracks very well when the vertical and horizontal lines are framed by a floating outline square, but his performance goes to pieces when the frame is removed so that the chopstick illusion reappears. The problem is that the sliding intersection is not parsed as an object. This suggests that pursuit eye movements are not driven by "dumb clockwork" but are instead dependent upon top-down signals based on object parsing. This almost 10-fold ratio in tracking errors suggests that smooth pursuit movements are not merely a bottom-up retinal feedback system (Krauzlis & Lisberger, 1994; Lisberger, Morris, & Tychsen, 1987) but may be strongly influenced by top-down cognitive processes such as object interpretation (Kowler, 1990).

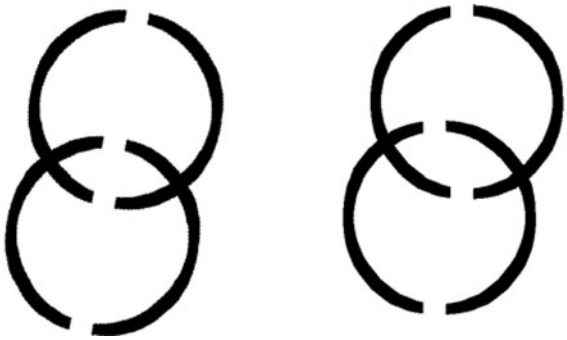
Sliding Rings

In Video IV.64-17 the terminators were removed by bending the lines around into rings. Two dots or gaps painted on each ring played the role previously played by the line terminators. These local dots radically altered the global percept of the entire rings. When the dots rotated in step with the rings (Video IV.64-17a), observers reported a single rigid welded figure eight, rotating coherently. This satisfies the rigidity constraint (Ullman, 1979). When the dots floated at 12 o'clock on each ring (Video IV.64-17b), observers reported two separate rings sliding over each other. This minimizes the motion within each ring by sacrificing rigidity (Shiffrar & Pavel, 1991; Ullman, 1979). Thus each ring coheres with its dot rather than with its intersection.

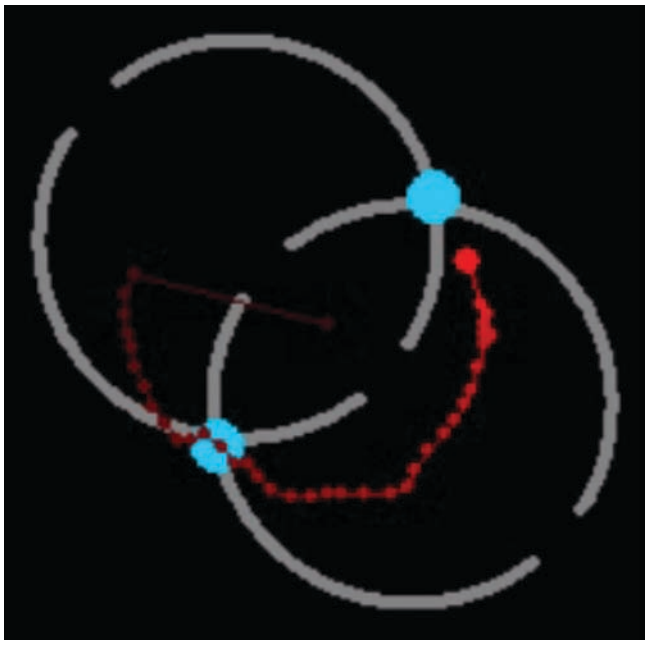
Video IV.64-18 shows that sliding rings caused the same collapse in eye movements that Video IV.64-16 produced with sliding lines. Observers had no trouble in tracing the apparently rigid rings from Video IV.64-17a, but they made gross tracking errors when viewing the apparently sliding



Video IV.64-16. Observers made pursuit eye movements to track first the rigid intersections in Video IV.64-12, right, and then the sliding intersections in Video IV.64-12, left. Tracking errors were 10 times greater in the latter case. (The blue spot was not present in the original tracked stimulus.) (Video by Hiroyuki Ito, Kyushu U.)



Video IV.64-17. Both rotating figure eights are identical. But (left) when the white gaps are painted on to the rings and rotate with them, observers report a single rigid figure eight rotating. When (right) the gaps "float" at 6 and 12 o'clock, the figure separates perceptually into two separate rings sliding over each other. (Video by Stuart Anstis.)



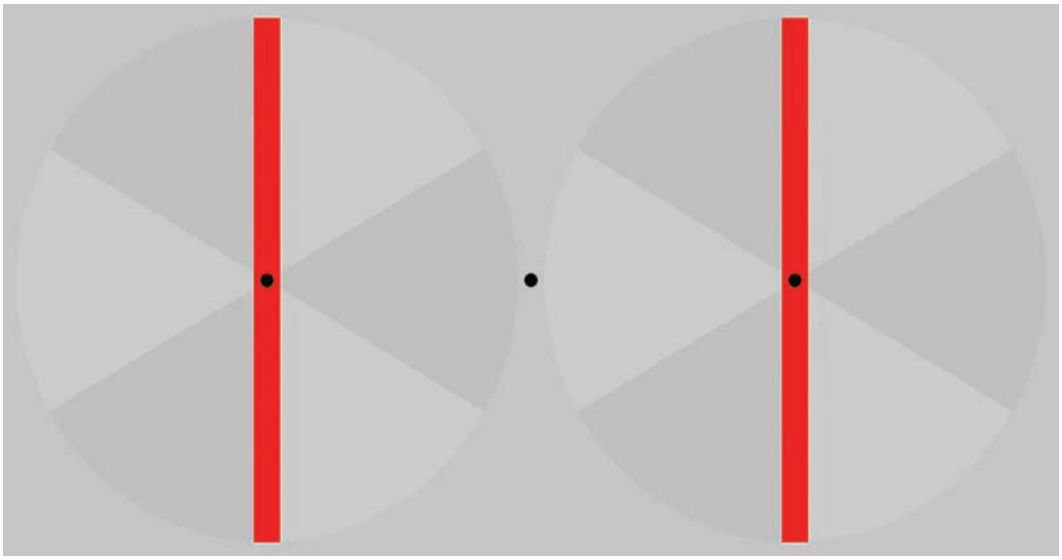
Video IV.64-18. Observers' eye movements pursued first the rigid intersections in Video IV.64-17, left, and then the sliding intersections in Video IV.64-17, right. Tracking errors were 10 times greater in the latter case. (The blue spot was not present in the original tracked stimulus.) (Video by Hiro Ito, Kyushu U.)

rings in Video IV.64-17b. Once again, the blue disks marking the intersections were not shown to the observer. These results cannot be predicted from the vague idea that the visual system prefers "simplicity" or a "good Gestalt."

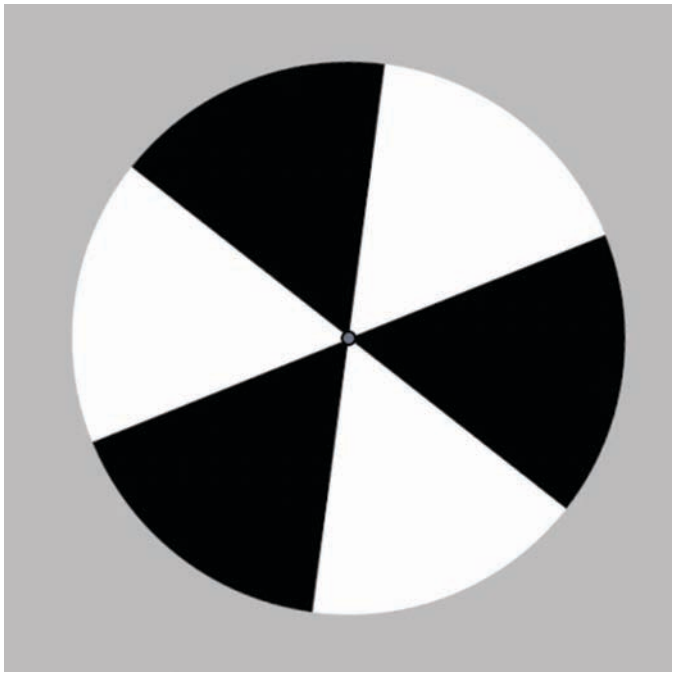
THE FLASH GRAB ILLUSION

The flash grab illusion is a new(ish) illusion in which the positions of lines are misperceived when they are flashed up on rotating backgrounds (Cavanagh & Anstis, 2013). The lines get apparently dragged along in the direction of the underlying movement.

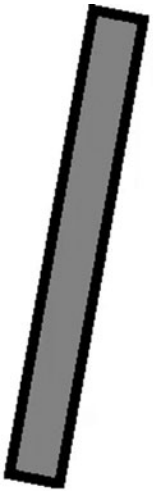
Look at Video IV.64-19. At first you see red vertical lines, flashing in alternation with green vertical lines. Note that the red and green lines are truly vertical and congruent. But when rotating sectored disks fade up into view behind them, the red lines appear to be tilted inward like the side bars of a capital A, while the green lines appear to tilt outward like the letter V. The effect is greatest when, as here, the lines are flashed up just as the direction of the rotating disks reverses. On the left-hand disk, the red line is dragged CW by the CW motion that comes just after the red flash, and similarly the green flashed line is dragged CCW by the CCW motion that follows it. Observers can look straight at either disk, but the effect is best if one gazes at the spot halfway between the two disks.



Video IV.64-19. Flashed red and green bars are always vertical. But when superimposed on sectored disks that rotate back and forth, the bars are perceptually sucked along in the direction of the motion that follows the flashes. They then appear to tilt back and forth. (Video by Patrick Cavanagh (Dartmouth) & Stuart Anstis.)



Video IV.64-20. Flash grab illusion: "Wonky cross." Same as Video IV.64-19 except with horizontal cross bars added. The crosses look tilted, and their intersections no longer look orthogonal. This is because the moving black/white edges of the rotating sectored disk grab the verticals and have more pulling power than the midpoint of sectors that grab the horizontals. (Video by Patrick Cavanagh (Dartmouth) & Stuart Anstis.)



Video IV.64-21. Flash grab illusion. Bar rotates back and forth through 180°. At slow speeds this is seen veridically, but at fast speeds the trajectory is underestimated and looks like ~120°, never quite reaching the vertical ("from 1 o'clock down to 5 o'clock"). This underestimation is thought to underlie the flash grab illusion. (Video by Patrick Cavanagh (Dartmouth) & Stuart Anstis.)

Video IV.64-20 (the "wonky cross") is similar, except that instead of just vertical lines an upright red cross now flashes up in alternation with a congruent upright green cross. Of course the two crosses are exactly superimposed. But when the rotating disk fades in, two things happen: First, the verticals are apparently tilted out of the vertical, as before. Second, the horizontals are tilted by different amounts, such that the arms of each cross appear to intersect at an acute angle instead of at a right angle. The reason is that the verticals of each cross lie on the edge of a rotating sector of the background disk, while the horizontal arms lie on the middle of rotating sectors. Clearly, the moving sector edges drag the arms of the cross much more effectively than the midsectors do.

What causes the flash grab effect? Basically it is caused by a perceptual shortening of motion trajectories (Sinico, Parovel, Casco, & Anstis, 2009). Video IV.64-21 shows a vertical bar that rotates slowly CW through 180° and back again. The top end of the bar is correctly seen to rotate back and forth between 12 o'clock and 6 o'clock. But when the bar speeds up, its trajectory is grossly underperceived and it seems rotate back and forth between about 1 o'clock and 5 o'clock. Suppose that the bar were to flash red each time it reached the vertical (not shown); the bar would appear to be tilted first toward 1 o'clock and then toward 5 o'clock. This is the flash grab effect. The flashes play no causal role in the illusion; they simply mark out the perceived endpoints of the rotating bar. To explain the apparent shortening of the motion trajectory, we conjecture that the visual system keeps a running average, with a time window of about 100 ms, of the position of the rotating bar. This averaging will round off the ends of the trajectory and make it appear shorter than it really is. Clearly the position of a moving object is not coded in some pure form but is influenced or contaminated by the context of its own motion. Although the flash grab effect most resembles the flash drag effect, it differs from this in the following ways: (a) it has a different temporal profile, (b) it requires attention, and (c) it is about 10 times larger.

REFERENCES

Anstis, S., & Ito, H. (2010). Eyes pursue moving objects, not retinal motion signals. *Perception*, *39*(10), 1408–1411.

Anstis, S., & Kim, J. (2011). Local versus global perception of ambiguous motion displays. *Journal of Vision*, *11*(3),13.

Anstis, S., & Ramachandran, V. S. (1986). Entrained path deflection in apparent motion. *Vision Research, 26*(10), 1731–1739.

Anstis, S., & Ramachandran, V. S. (1987) Visual inertia in apparent motion. *Vision Research, 27*(5), 755–764.

Bruno, J., & Bertamini, M. (2013). Perceptual organization and the aperture problem. In J. Wagemans (Ed.), *The Oxford handbook of perceptual organization* (pp. 504–520). Oxford: Oxford University Press.

Burt, P., & Sperling, G. (1981). Time, distance, and feature trade-offs in visual apparent motion. *Psychological Review, 88*(2), 171–195.

Cavanagh. P., & Anstis, S. (2013). The flash grab effect. *Vision Research, 91*, 8–20.

Duncker, K. (1955). Induced motion. In W. D. Ellis (Ed.), *A source book of Gestalt psychology* (pp. 161–172). London: Routledge & Paul. (Original work published 1929)

Farrell-Whelan, M., Wenderoth, P., & Wiese, M. (2012). Studies of the angular function of a Duncker-type induced motion illusion. *Perception*, *41*(6), 733–746.

Gengerelli, J. A. (1948). Apparent movement in relation to homonymous and heteronymous stimulation of the cerebral hemispheres. *Journal of Experimental Psychology*, *38*(5), 592–599.

Kowler, E. (1990). *Eye movements and their role in visual and cognitive processes*. New York: Elsevier.

Krauzlis, R. J. & Lisberger, S. G. (1994). A model of visually-guided smooth pursuit eye movements based on behavioral observations. *Journal of Computational Neuroscience*, *1*, 265-83.

Liden, L. & Pack, C. (1999). The role of terminators and occlusion cues in motion integration and segmentation: a neural network model. *Vision Research*, *39*, 3301-20.

Lisberger, S. G., Morris, E. J., & Tychsen, L. (1987). Visual motion processing and sensory-motor integration for smooth pursuit eye movements. *Annual Review of Neuroscience*, *10*, 97–129.

Metelli, F. (1974). The perception of transparency. *Scientific American*, *230*(4), 90–98.

Pinkus, A., & Pantle, A. (1997). Probing visual motion signals with a priming paradigm. *Vision Research, 37*(5), 541–552.

Ramachandran, V. S., & Anstis, S. M. (1983). Perceptual organization in moving patterns. *Nature*, *304*(5926), 529–531.

Shiffrar, M., & Pavel, M. (1991). Percepts of rigid motion within and across apertures. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(3), 749–761.

Shimojo, S., Silverman G. H., & Nakayama, K. (1989). Occlusion and the solution to the aperture problem for motion. *Vision Research*, *29*, 619–626.

Sinico, M., Parovel, G., Casco, C., & Anstis, S. (2009). Perceived shrinkage of motion paths. *Journal of Experimental Psychology. Human Perception and Performance*, *35*(4), 948–957.

Stoner, G. R., & Albright, T. D. (1994). Visual motion integration: A neurophysiological and psychohysical perspective. In A. T. Smith & R. J. Snowden (Eds.), *Visual detection of motion* (pp. 279–280). London: Academic Press.

Ullman, S. (1979). The interpretation of structure from motion. *Proceedings of the Royal Society of London, B*, *203*(1153), 405–426.

Zivotofsky, A. Z. (2004). The Duncker illusion: Intersubject variability, brief exposure, and the role of eye movements in its generation. *Investigative Ophthalmology*, *45*, 2867–2872.

Zivotofsky, A. Z., Goldberg, M. E., & Powell, K. D. (2005). Rhesus monkeys behave as if they perceive the Duncker illusion. *Journal of Cognitive Neuroscience, 17*(7), 1011–1017.

Chapter 65
Backscroll Illusion

Kiyoshi Fujimoto

THE PHENOMENON

The backscroll illusion is seen in a video clip that presents a walker or a runner against a background of a counterphase grating pattern (Fig. IV.65-1). The overall image stands at one position on a screen. The walker/runner looks like he is on a treadmill. The grating has just enough periodic reversals of luminance contrast that it is directionally ambiguous. However, the grating appears to scroll backward to the walker/runner. The same effect is produced from another mobile object such as an animal or a vehicle.

EMPIRICAL RESULTS

Fujimoto and Sato (2006) conducted psychophysical experiments to show that recognition of object motion implied by bodily action and/or form determines the backscroll illusion. In the first experiment, the authors manipulated velocities of the walker/runner and the grating. A walking/running velocity was defined by the speed of pendulum-like movements of the arms and legs. The counterphase grating was actually composed of two gratings moving in the opposite directions at the same velocity. The velocity of the component gratings could be controlled by keeping the appearance of the overall grating directionless. As a result, the more similar the velocities of the walker/runner and the grating were, the more often the illusion occurred. Next the authors investigated a time course and found that the illusion arose as if it was synchronized with recognition of bodily action. In addition, the illusion sustained against several reversals of limb swings. It denied physical movements of the limbs in a specific period and/or area as a determinant. In the third experiment, the authors showed that even a still picture of the walker induced the illusion, although its probability was about half of that yielded by the animated walker.

These experiments also clarified optimal conditions of the backscroll illusion. First, the apparent walking velocity is similar to the component velocity of the counterphase grating. Second, the grating has a temporal frequency of 10 to 20 Hz. Third, duration of the video is 0.2 to 4 s. Finally, the human figure has bodily movements of forward walking. As another factor, from my personal impression, peripheral vision offers a clearer illusion than central vision, although this has not been found in published experiments (Fujimoto & Yagi, 2008).

Fujimoto, Yagi, and Sato (2009) quantified strength of the backscroll illusion under optimal conditions using a psychophysical motion-nulling method. The counterphase grating has unidirectional appearance when the component gratings have different luminance contrasts. Perceived motion is dominant in the direction of the component grating with a higher contrast. The authors used this characteristic to determine a contrast ratio where the walking figure canceled the unidirectional perception in the grating. Results showed that the backscroll illusion disappeared when a luminance contrast of the component grating moving in the walker's direction was about twice as high as that for the other grating. This means that the walker figure had a bias comparable to the physical motion energy of the component grating.

Fujimoto, Yagi, and Sato (2009) reanalyzed data of the motion-mulling experiment to investigate variability between and within individuals. Large individual differences were suggested across 50 observers. The most sensitive observer perceived the illusion above five times as strongly as the least sensitive observer in terms the nulling contrast ratio. On the other hand, the ratios of the maximum to minimum nulling ratios from three measurements for 17 observers were very low, indicating small variability within individuals.

Fujimoto and Yagi (2008) showed that another kind of stimulus produces the backscroll illusion. The authors presented a walker as a biological motion figure in which only small point-lights corresponding to the main joints of limbs were visible. The background was a random-dot cinematogram that consisted of some signal dots moving coherently either leftward or rightward among other noise dots moving randomly. Increasing the proportion of the signal dots enhanced perception of coherent motion of the signal dots. The authors changed the proportions randomly trial by trial, and observers judged a perceived direction of coherent motion. As a result, the point-light walking figure produced coherent motion perception opposite to the walking direction with a smaller proportion of the signal dots than a control figure that was not recognized as a



Figure IV.65-1. A snapshot of a video clip producing the backscroll illusion.

walker (accomplished by scrambling the configurations of the point-lights). Fujimoto (2003) reported the backscroll illusion from a combination of the point-light walker and the counterphase grating. Generally, the backscroll illusion from the biological motion figure is weak. Also, from my personal impression, the illusion in the random-dot cinematogram is weak even using the realistic human figure.

THEORETICAL INTERESTS

The backscroll illusion is unique in that implied object motion causes apparent motion in background or adjacent stimuli. As a similar illusion, one may remember an induced motion phenomenon in which a static or ambiguously moving stimulus appears to move in the opposite direction of nearby unidirectional stimuli (see Reinhardt-Rutland, 1988, for a review). The induced motion illusion is well explained by center-surround antagonism in the integration of local motion signals (see Nishida, 2011, for a review). Motion signals at one position receive oppositely directed modulation from surrounding areas. Representing real motion is a general function for the visual system, so it also works in the processing of the stimulus causing the backscroll illusion. However, it never explains the final percepts because there is no bias in a total of real motion components (Fujimoto & Sato, 2006). It is better understood by the fact that the static walker figure produces the illusion. Therefore, the backscroll illusion needs a new explanation.

Even a still picture implies movements of a familiar object, as we can recognize the person walking leftward in Figure IV.65-1. Also, such a picture produces sensation of motion (see Cutting, 2002, for a review) and activates brain areas involved in processing real motion (see Kourtzi, Krekelberg, & van Wezel, 2008, for a review). A psychophysical study showed that viewing pictures of moving people or animals produced a motion aftereffect in a real motion display toward the opposite direction of living things (Winawer, Huk, & Boroditsky, 2008). Although this aftereffect is similar to the backscroll illusion, it needs a different explanation. The aftereffect was most likely caused by activation of real motion mechanisms to represent the movements of objects themselves. On the other hand, in the backscroll illusion, it is the background that reflects an influence of implied motion on real motion mechanisms. In addition, the backscroll illusion clearly demonstrates a simultaneous effect of implied motion on perception of real motion. Thus it is plausible that modulation of real motion signals by implied motion signals underlies the backscroll illusion.

Why does implied motion modulate real motion? One simple answer is because of learning. In everyday environments, we often see moving objects in front of static backgrounds. For example, in town, pedestrians or cars travel in front of houses, buildings, or walls. In natural scenery, animals run in front of bushes, trees, or mountains. Eye tracking to the object in such situations creates a retinal image corresponding to the backscroll illusion. A neurophysiological study showed that motion-sensitive neurons of macaque monkeys respond to a static picture after associative learning with a real motion stimulus (Schlack & Albright, 2007). However, the neurons activated 0.5 s from stimulus onset, which is inconsistent with the psychophysical result showing that the backscroll illusion reaches its maximum before 0.5 s

(Fujimoto & Sato, 2006). This discrepancy may depend on experimental stimuli or living environments of monkeys and humans. Further investigations are needed to make a conclusion.

Finally, it can be hypothesized that the backscroll illusion reflects a function that implied object motion serves as a frame of reference to disambiguate or recalibrate retinal flow signals. Retinal motion signals are inherently ambiguous due to aperture problem (see Nishida, 2011 for a review) and observers' movements. On the other hand, several studies have confirmed that perception of biological motion is robust against various noises (see Fujimoto & Sato, 2006, for a review). Tadin, Lappin, Blake, and Grossman (2002) showed that the human form defined by biological motion serves as a frame of reference for the evaluation of local motion coherency. Seno and Sato (2012) showed that viewing walkers induces vection, namely, sensation of self-motion. Therefore it is possible that the visual system uses object motion cues to stabilize the perceptual world. However, this hypothesis also requires further investigation.

REFERENCES

Cutting, J. E. (2002). Representing motion in a static image: Constraints and parallels in art, science, and popular culture. *Perception*, *31*, 1165–1193.

Fujimoto, K. (2003). Motion induction from biological motion. *Perception*, *32*, 1273–1277.

Fujimoto, K., & Sato, T. (2006). Backscroll illusion: Apparent motion in the background of locomotive objects. *Vision Research*, *46*, 14–25.

Fujimoto, K., & Yagi, A. (2007). Backscroll illusion in far peripheral vision. *Journal of Vision*, *7*(8):16, 1–7.

Fujimoto, K., & Yagi, A. (2008). Biological motion alters coherent motion perception. *Perception*, *37*, 1783–1789.

Fujimoto, K., Yagi, A., & Sato, T. (2009). Strength and variability of the backscroll illusion. *Vision Research*, *49*, 759–764.

Kourtzi, Z., Krekelberg, B., & van Wezel, R. J. A. (2008). Linking form and motion in the primate brain. *Trends in Cognitive Neuroscience*, *12*, 230–236.

Nishida, S. (2011). Advancement of motion psychophysics: Review 2001–2010. *Journal of Vision*, *11*(5):11, 1–53.

Reinhardt-Rutland, A. H. (1988). Induced movement in the visual modality: An overview. *Psychological Bulletin*, *103*, 57–71.

Schlack, A., & Albright, T. D. (2007). Remembering visual motion: Neural correlates of associative plasticity and motion recall in cortical area MT. *Neuron*, *53*, 881–890.

Seno, T., & Sato, T. (2012). Vection can be induced without explicit motion signal using backscroll illusion. *Japanese Psychological Research*, *54*, 218–222.

Tadin, D., Lappin, J. S., Blake, R., & Grossman, E. D. (2002). What constitutes as efficient reference frame for vision? *Nature Neuroscience*, *5*, 1010–1015.

Winawer, J., Huk, A. C., & Boroditsky, L. (2008). A motion aftereffect from still photographs depicting motion. *Psychological Science*, *19*, 276–283.

Chapter 66
The Rotating Tilted Lines Illusion

Rotation Instead of Expansion—A Peculiar Case of Motion Misperception

Simone Gori

INTRODUCTION

"Motion illusion" refers to a perception of motion that is absent and/or different (in direction, strength, etc.) in the physical stimulus but is still perceived by the observer (see Gori & Stubbs, 2014 for a review). The illusory motion is experienced as "real motion," and it should not be confused with the perception of dynamism evocated by paintings (Gori, Pedersini, & Giora, 2008) or by the speed lines often used in comics (Burr, 2000). Motion illusions are a peculiar case of visual illusions that, because of their striking feature, have captured the attention of vision scientists but also of artists, thus reaching a broader audience than is normally captivated by a scientific topic. There are many possible ways to classify motion illusions (see Burr & Thompson, 2012; Ito, 2012, Kitaoka & Ashida, 2007, for reviews). One option is to divide them into two categories based on the type of their illusory motion: (a) the illusory motion observed without any type of motion in the stimulus and (b) the misperception of motion, that is, the perception of motion that differs from the physical motion of the stimulus. Both types of illusions produce excellent examples of motion illusions that provide knowledge regarding the underlying visual system mechanisms (Spillmann, 2009).

A few examples of the first category are the MacKay rays (MacKay, 1957), the Enigma illusion (i.e., Gori, Hamburger, & Spillmann, 2006; Gregory, 1993; Hamburger, 2007; Kumar & Glaser, 2006; Leviant, 1996, Mon-Williams & Wann, 1996; Ruzzoli et al., 2011; Troncoso, Macknik, Otero-Millan, & Martinez-Conde, 2008; Zeki, Watson, & Frackowiak, 1993; see chapter 67 in this volume for details), the Ouchi illusion (i.e., Fermüller, Pless, & Aloimonos, 2000; Mather, 2000; Spilmann, 2013; Spillmann, Heitger, & Schuller, 1986; Spillmann & Pinna, 2005; Spillmann & Werner, 1990), and the Fraser-Wilcox/rotating snakes illusion (i.e., Agrillo, Gori & Beran, 2015; Backus & Oruç 2005; Beer, Heckel, & Greenlee, 2008; Conway, Kitaoka, Yazdanbakhsh, Pack, & Livingstone, 2005; Faubert & Herbert, 1999; Fraser & Wilcox, 1979; Gori, Agrillo, Dadda & Bisazza, 2014; Kitaoka & Ashida, 2003; Kuriki, Ashida, Murakami, & Kitaoka, 2008; Murakami, Kitaoka, & Ashida, 2006; Naor-Raz & Sekuler, 2000; Otero-Millan, Macknik, & Martinez-Conde, 2012; see chapter 68 in this volume for details).

The second category of motion illusions, the misperception of motion, is characterized by the presence of physical motion that is incorrectly processed by the visual system that ends up as a percept that differs from the original motion. The illusion that is the subject of this chapter falls into this category and represents the basis of more complex patterns. A few examples of this category are the barber pole illusion (i.e., Guilford, 1929; Lidén & Mingolla, 1998; Wallach, 1935)—which also was the basis of similar phenomena described many years later such as, for example, the oblique-line illusion (Bressan & Vezzani, 1995) or the nonrigid illusory motion showed by Kim and Papathomas (2011); the revolving wheels illusion, also known as the Pinna-Brelstaff illusion (i.e., Gurnsey & Pagé, 2006; Gurnsey, Sally, Potechin, & Mancini, 2002; Pinna & Brelstaff, 2000; see chapter 74 in this volume for details); the unchained dots illusion (Gori, Giora, & Stubbs, 2010); and the accordion grating illusion (Gori, Giora, Yazdanbakhsh, & Mingolla, 2011; Giora, Yazdanbakhsh, & Mingolla, 2013; Yazdanbakhsh & Gori, 2011).

DESCRIPTION OF THE PHENOMENON

The phenomenon reported in this chapter is characterized by misperception of motion and is probably the simplest pattern able to elicit rotary motion in the presence of physical radial motion. Figure IV.66-1 represents the rotating-tilted-lines illusion (RTLI; Gori & Hamburger, 2006). In its simplest form (Fig. IV.66-1a), this new illusion arises in a circular pattern of 120 black radial lines tilted to the right and presented on a white background. When one approaches the stimulus pattern, the radial lines appear to rotate in a counterclockwise direction, whereas when one recedes from it, they appear to rotate clockwise. If the lines are tilted to the left, the illusory effects occur in the opposite direction as a function of the self-motion. Gori and Hamburger also created a more complex version of this illusion (from a strict chronological point of view it was the first one created by the authors). The ring of tilted lines was embedded in a context of an equal number of lines

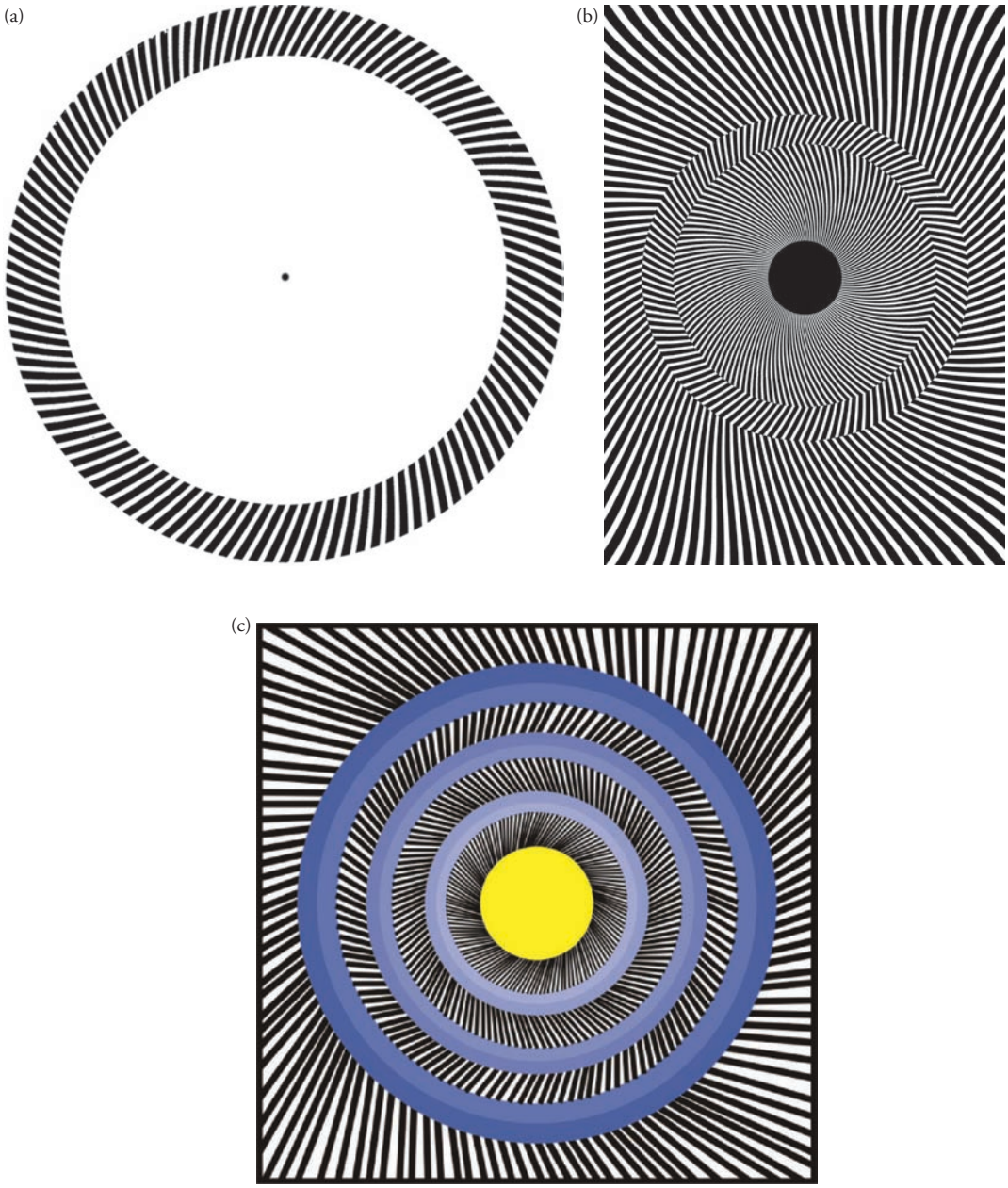


Figure IV.66-1. Panel (a): The simplest version of motion illusion that requires self-motion. One hundred twenty black radial lines are tilted to one side. When one approaches the stimulus pattern, the radial lines appear to rotate in a counterclockwise direction, whereas when one recedes from it, they appear to rotate clockwise. Panel (b): A more complex version of the previous illusion; in this pattern the illusory rotation is stronger and there may be some residual counter-rotation in the surround. Panel (c): The illusory motions of the RTLI and the Enigma illusion become united in a single pattern. Approaching and receding from the pattern leads to rotation of the tilted radial lines in clockwise or counterclockwise direction, while fixating the center of the static patterns leads to spinning motion on the three colored rings. (Figures created by Simone Gori.)

oriented in the opposite direction representing the background (Fig. IV.66-1b). Here the illusory rotation is stronger, and there may occur some residual counterrotation in the surround. Hine, Cook, and Rogers (1995, 1997) showed similar patterns reporting only the slide motion (similar to the Ouchi illusion) and not the circular motion due to the observer motion, which is the main effect of the RTLI. Gori and Hamburger also created a pattern able to elicit two different kinds of illusory motion: a spontaneous one and one due to self-motion. This pattern with the two illusory effects combined is depicted in Figure IV.66-1c. The tilted line pattern of Figure IV.66-1a was mixed up with the rings of the Enigma illusion (Leviant, 1996). The effect of perceived clockwise and counterclockwise rotation is still present on approaching the pattern and receding from it. Moreover, in static view, streaming motion on the rings appears as it does in the Enigma illusion.

Interestingly, Billino, Hamburger, and Gegenfurtner (2009) investigated the age effect on several motion illusions, and the RTLI was the only one among the tested

illusions to be perceived as often by children (3 to 6 years) as by young adults (18 to 39 years): 95.8% of the children and 95.6% of the young adults reported illusory rotation. Surprisingly, the revolving wheels illusion (Pinna & Brelstaff, 2000), which shares some characteristics with the RTLI, was perceived by only 16.7% of the children, suggesting the RTLI as the best pattern to study illusory motion at an early age. Moreover, according to Billino et al. (2009), the illusory rotation of the RTLI is perceived more often than the one elicited by the revolving wheels during the age period from 3 to 64 years, making the RTLI not only the simplest pattern able to elicit illusory rotation due to the observer motion but also the more powerful. Finally, it is worth noting that the RTLI breaks down (or it is a least strongly reduced) if presented in isoluminant colors (Hamburger, 2012). This results shows that this illusion is probably processed almost exclusively by the Magnocellular-Dorsal pathway which is blind to colors. Recently, the RTLI was used as a tool to investigate the Magnocellular-Dorsal pathway functionality in children with developmental dyslexia (Gori et al., 2015; Gori, Seitz, Ronconi, Franceschini, & Facoetti, 2016; Gori, Molteni, & Facoetti, 2016).

DISCUSSION

Orientation of the lines seems to be the main factor in eliciting the illusory motion in the RTLI (Gori & Hamburger, 2006; Gori & Yazdanbakhsh, 2008). Gurnsey et al. (2002) highlighted how also in the Pinna-Brelstaff illusion (Pinna & Brelstaff, 2000) the low-frequency orientation of the micropatterns is crucial for the illusory motion and suggested direction-selective neurons as a neutral substrate of this perceived rotation. They reduced the Pinna-Brelstaff pattern to a circle of Gabor patches; no ends of the lines were left, which perfectly served the purpose of their study. Unlike this, the RTLI is composed of simple lines with extra cues of line ends, which aids correct motion detection compared with the ambiguous motion signal in the middle of the lines. The origin of the RTLI illusory phenomenon may be understood as a consequence of the coding of ambiguous and nonambiguous motion signals afforded by the stimulus (Gurnsey et al., 2002; Yazdanbakhsh & Gori, 2008) or, equivalently, as a result of competition between motion signals originating from two different classes of motion-processing units (Gori et al., 2011; Gori & Yazdanbakhsh, 2008; Grossberg & Mingolla, 1993; Lorenceau, Shiffrar, Wells, & Castet, 1993). Lorenceau et al. investigated a systematic error in direction discrimination of lines moving obliquely relative to their orientation. Two different units in the early visual cortex are supposed to be involved in this misperception of direction discrimination. Some motion-processing units maximally respond when the full length of a line passes over their receptive field (called contour units). These units face the aperture problem: they can only pick up the normal component of velocity from any straight line or edge that fills their receptive field (Adelson & Movshon, 1982; Fennema & Thompson, 1979). The classical aperture problem confronted by cells with spatially limited receptive fields can be viewed as follows: each neuron in the visual system is sensitive to an input in a small portion (the aperture) of the visual field, as if each neuron is looking through a window. The motion direction of a contour is ambiguous, because the motion component parallel to the line cannot be measured based on the visual input. This means that a variety of contours with different orientations moving at different speeds can cause identical responses in a motion-sensitive neuron (Gori et al., 2011).

The second type of motion-processing unit responds to a moving line best when one or other line ends (terminators) fall within their receptive field. Such units, symmetrical or asymmetrical end-stopped neurons, can indicate the true direction of motion of the entire line from information afforded by the line-end motion (Gori et al., 2011; Mingolla, Todd, & Norman, 1992; Pack, Livingstone, Duffy, & Born, 2003). An idealization of these two types of motion-processing units appears in Figure IV.66-2. The subjects' misperception of line directions could happen because the overall response of the contour units overcomes the response to terminator motion. Gori and Yazdanbakhsh (2008) proposed a model for the RTLI based on these two types of motion-processing units that can be thought of as a key functional component "extracted" from more complex motion-processing models (Bayerl & Neumann, 2004; Berzhanskaya, Grossberg, & Mingolla, 2007; Chey, Grossberg, & Mingolla, 1997; Grossberg & Mingolla, 1993; Grossberg, Mingolla, & Viswanathan, 2001). The assumption of two units was borrowed from the findings of single-cell electrophysiology and offers an attempt to reconcile the existence of different units and the psychophysical percept. Approaching the RTLI, we have an expansion motion of the circle on the retina (Fig. IV.66-3). The motion of every single line composing the circle is analyzed by n contour units and n end-stopped units. The end-stopped units signal the correct direction of motion (Fig. IV.66-2a; solid black arrows in Fig. IV.66-3), while the n contour units signal a direction that is different from the veridical motion signal available at the line endings (Fig. IV.66-2b; dashed gray arrows in Fig. IV.66-3). The ratio between the overall response of end-stopped units and contour units determines the final signaled direction. In a simplified formalism, consider the end-stopped and contour units E_i^k and C_i^k that are tuned to motion direction k at location i along a particular line (Fig. IV.66-2). The overall activity of each set of E and C units for each direction is determined by

$$E^k = \sum_{i=1}^{N} E_i^k \tag{1}$$

$$C^k = \sum_{i=1}^{N} C_i^k \tag{2}$$

The net result of such calculations for an expanding tilted line is shown in Figure IV.66-2. As can be seen in Figure IV.66-2a, at the end of the lines (units 1 and N) it is the end-stopped cells that are the active units, while the end-stopped units within the line do not contribute to the summation in Equation 1 because end-stopped units do not respond to contours that fill their receptive field. The contour units instead receive their optimal input from the interior of the line, away from the line ends. Due to their

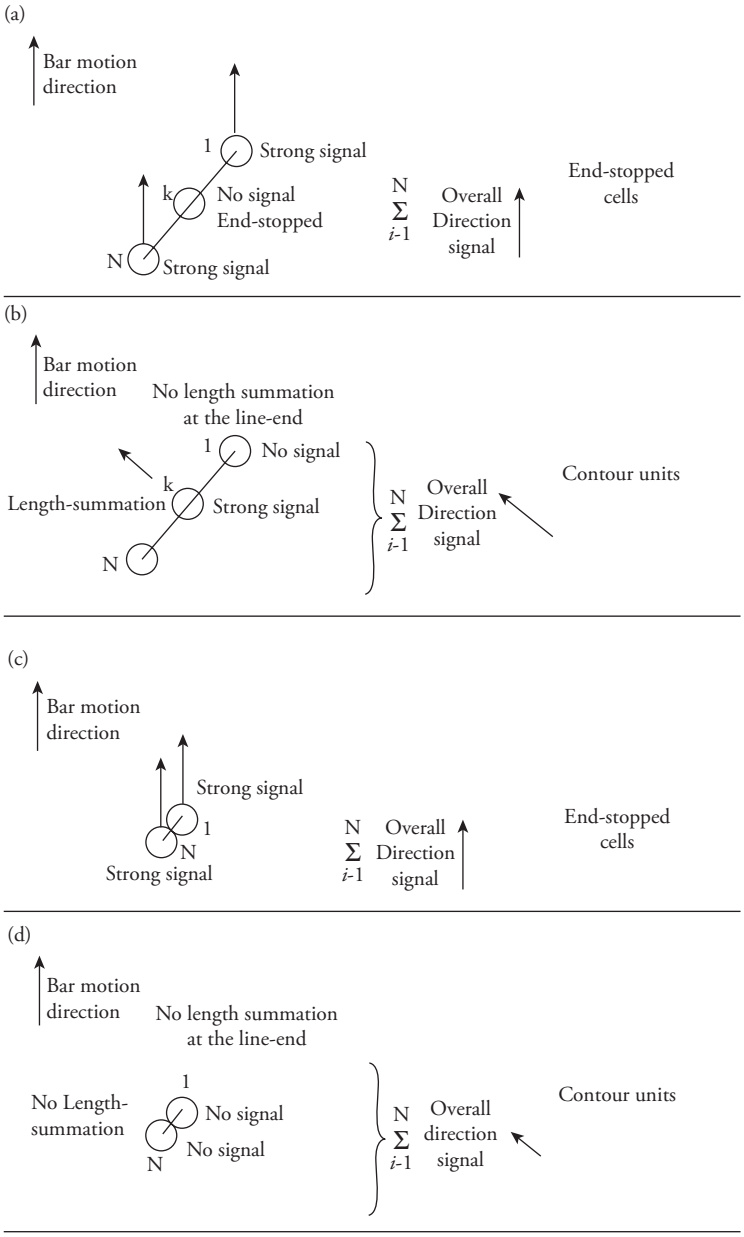


Figure IV.66-2. Panels (a) and (b): Idealized depiction of the two types of motion processing units that are proposed to be involved in the RTLI illusory rotation. Panels (c) and (d): The same units processing a very short tilted line moving upward. (Figures created by Simone Gori and Arash Yazdanbakhsh.)

length summating nature, they receive suboptimal input at the line ends because their receptive fields are not filled by the line. The net result for Equation 1 is to signal the veridical motion, while the result of Equation 2 is to signal orthogonal to the line motion.

In summary, the net result of Equation 1 would be that shown in Figure IV.66-2a together with that shown in Figure IV.66-2b. The weighted sums of end-stopped and contour units then determine the final direction of motion. The average would be weighted more toward the contour units than the end-stopped units, given that the length of the line is large enough to weigh the sum in Equation 2 above that in Equation 1. Given the schema depicted in Figure IV.66-2 and the involvement of the end units in the end-stopped group to signal the correct motion of the line, the only chance for this group to bias the weighted average more toward itself would be to use short line elements, because in short lines the activities of the 1st and *N*th unit in the end-stopped group are balanced by the lack of activities of the middle contour units, which do not produce any input (Figures IV.66-2c and IV.66-2d). According to Lorenceau et al. (1993), the line length should have an effect on the strength of the illusory motion. With very short lines the illusory rotation disappears (Gori & Hamburger 2006; Yazdanbakhsh & Gori, 2008). This is interpretable as a case where the activity of the contour units is not sufficient to bias the weighted activity toward them compared with the end-stopped units (Figures IV.66-2c and IV.66-2d). Based on that hypothesis, Yazdanbakhsh and Gori (2008) were able to measured the receptive fields size psychophysically, matching the previous electrophysiological results. The radii and tilted line length of RTLI were systematically changed while the subjects rated the illusory strength. The outcome of this

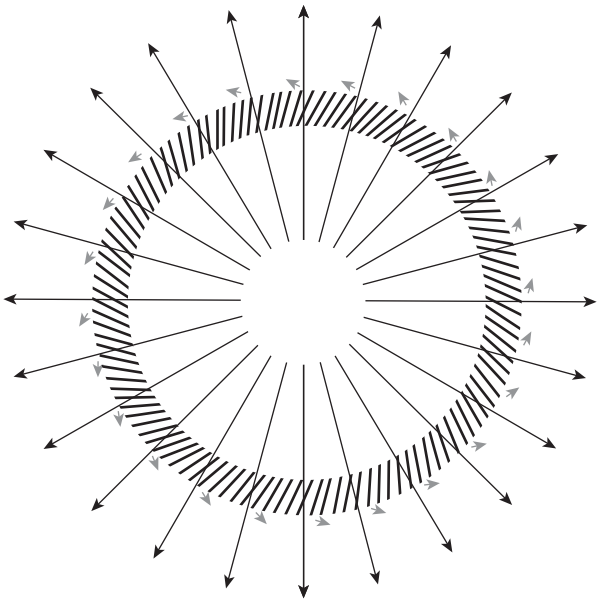


Figure IV.66-3. Expansion motion of the circle on the retina. The correct direction of motion is analyzed by end-stopped units (black arrows) while the contour units (dashed gray arrows) signal a direction different from the veridical motion signal. (Figure created by Simone Gori and Arash Yazdanbakhsh.)

experiment provided a very accurate measurement of the receptive fields size at different eccentricity and indirectly confirmed the Gori and Yazdanbakhsh (2008) model. Other observations supported the Gori and Yazdanbakhsh model as mentioned later. Inverting the orientation of the lines in the RTLI results in an inversion of illusory rotation (Gori & Hamburger, 2006). This is in accord with the Gori and Yazdanbakhsh model, because the motion direction signaled by the contour units is also inverted. The illusion disappears if the lines are arranged radially. This result is expected by the proposed model because all of the motion-processing units will signal the same true direction of motion. Moreover, an RTLI composed of dashed lines reduces drastically the illusory motion (see Gori & Yazdanbakhsh, 2008, Fig. 5). This can be derived from the same explanation: The very short stripes weight the percept toward the motions signaled by end-stopped units more than those signaled by the contour units. Finally, to confirm the involvement of the end-stopped neurons Gori and Yazdanbakhsh created and empirically tested the pattern in Figure IV.66-4. This pattern generates almost no signal for the end-stopped neurons due to the blurred edges. Only the contour units would be involved in the production of the final percept. The significant enhancement in illusory motion strength observed with this pattern supports the involvement of end-stopped neurons in the typical RTLI pattern.

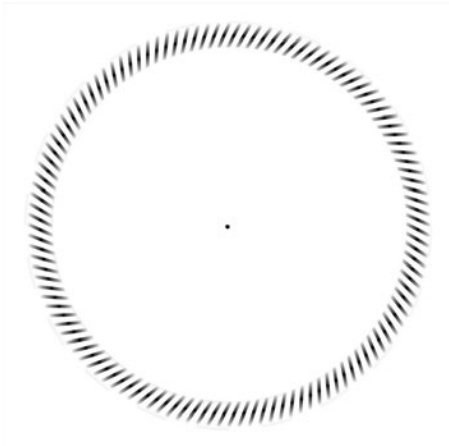


Figure IV.66-4. The RTLI produced by short lines with blurred edges. Approaching and receding from the pattern should produce a stronger rotation than the one experienced with Figure IV.66-1a. (Figure created by Simone Gori.)

Interestingly, in the aforementioned motion-processing models (Bayerl & Neumann, 2004; Berzhanskaya et al., 2007; Chey et al., 1997; Grossberg & Mingolla, 1993; Grossberg et al., 2001), the output of end-stopped units is often dispositive for the final model percept. That is, when processing a moving line segment, the signals from numerically fewer end-stopped units usually overcome the numerical superiority of aperture-ambiguous signals from contour units, thereby allowing for, say, veridical perception of rightward motion for a line that happens to be diagonally oriented in the visual field. The mechanisms for accomplishing this "undemocratic" feat include the down-regulation of aperture-ambiguous signals along the interiors of lines by shunting cross-direction competition among contour units. This was anticipated from psychophysics by Lorenceau et al. (1993) and Mingolla et al. (1992), first computationally simulated by Chey et al. (1997), and confirmed by single-unit electrophysiology by Pack and Born (2001); however, this process of global resolution of the aperture problem takes time (on the order of 100 ms.). As showed by Yazdanbakhsh and Gori, the simple voluntary regulation of the temporal frequency of expansion/contraction speed of the RTLI set the balance between the effects of end-stopped units and contour units. In other words, the illusory rotation can be reduced or even eliminated by simply slowing down the expansion/contraction speed. The result for variations in speed is consistent with the time required for integration processes (propagation of the end-stopped neuron signals and/or larger receptive fields in high cortical area contribution) during observation of the RTLI, as reported for MT neurons by Pack and Born in macaque monkeys, inferred psychophysically for humans by Lorenceau et al., and confirmed with the accordion grating nulling experiments by Gori et al. (2011).

In sum, this chapter reports a new illusory phenomenon, the RTLI, which seems to involve two motion-processing units represented by two families of motion-sensitive neurons in the early visual cortical areas. The RTLI presents a striking rotation in the presence of expansion/contraction motion on the retina despite its very simple pattern. The line orientation composing the pattern seems to be the main variable in producing the illusory effects.

REFERENCES

Adelson, E. H., & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, *300*(5892), 523–525.

Agrillo, C., Gori, S., & Beran, M.J. (2015). "Do rhesus monkeys (Macaca mulatta) perceive illusory motion?" *Animal Cognition*, *18*(4), 895–910.

Backus, B. T., & Oruç, I. (2005). Illusory motion from change over time in the response to contrast and luminance. *Journal of Vision*, *5*(11), 1055–1069.

Bayerl, P., & Neumann, H. (2004). Disambiguating visual motion through contextual feedback modulation. *Neural Computation*, *16*(10), 2041–2066.

Beer, A. L., Heckel, A. H., & Greenlee, M. W. (2008). A motion illusion reveals mechanisms of perceptual stabilization. *PLoS ONE*, *3*(7), e2741.

Berzhanskaya, J., Grossberg, S., & Mingolla, E. (2007). Laminar cortical dynamics of visual form and motion interactions during coherent object motion perception. *Spatial Vision*, *20*(4), 337–395.

Billino, J., Hamburger, K., & Gegenfurtner, K. R. (2009). Age effects on the perception of motion illusions. *Perception*, *38*(4), 508–521.

Bressan, P., & Vezzani, S. (1995). A new motion illusion related to the aperture problem. *Perception*, *24*(10), 1165–1176.

Burr, D. (2000). Are "speed lines: used in human visual motion? *Current Biology*, *10*(12), R440–R443.

Burr, D., & Thompson, P. (2011). Motion psychophysics: 1985–2010. *Vision Research*, *51*(13), 1431–1456.

Chey, J., Grossberg, S., & Mingolla, E. (1997). Neural dynamics of motion grouping: From aperture ambiguity to object speed and direction. *Journal of the Optical Society of America A*, *14*(10), 2570–2594.

Conway, B. R., Kitaoka, A., Yazdanbakhsh, A., Pack, C. C., & Livingstone, M. S. (2005). Neural basis for a powerful static motion illusion. *Journal of Neuroscience*, *25*(23), 5651–5656.

Faubert, J., & Herbert, A. M. (1999). The peripheral drift illusion: A motion illusion in the visual periphery. *Perception*, *28*(5), 617–621.

Fennema, C. L., & Thompson, W. B. (1979). Velocity determination in scenes containing several moving objects. *Computational Graphics and Image Processing*, *9*(4), 301–315.

Fermüller, C., Pless, R., & Aloimonos, Y. (2000). The Ouchi illusion as an artifact of biased flow estimation. *Vision Research*, *40*(1), 77–96.

Fraser, A., & Wilcox, K. J. (1979). Perception of illusory movement. *Nature*, *281*(5732), 565–566.

Gregory, R. L. (1993). A comment: Mackay rays shimmer due to accommodation changes. *Proceedings of the Royal Society of London, B*, *253*(1336), 123.

Gori, S., Agrillo, C., Dadda, M., & Bisazza, A. (2014). "Do fish perceive illusory motion?" *Scientific Reports*, 4:6443.

Gori, S., Giora, E., & Stubbs, D. A. (2010). Perceptual compromise between apparent and veridical motion indices: The unchained-dots illusion. *Perception*, *39*(6), 863–866.

Gori, S., Giora, E., Yazdanbakhsh, A., & Mingolla, E. (2011). A new motion illusion based on competition between two kinds of motion processing units: The accordion grating. *Neural Networks*, *24*(10), 1082–1092.

Gori, S., Giora, E., Yazdanbakhsh, A. & Mingolla, E. (2013). The novelty of the accordion grating illusion. *Neural Networks*, *39*, 52.

Gori, S., & Hamburger, K. (2006). A new motion illusion: The rotating-tilted-lines illusion. *Perception*, *35*(6), 853–857.

Gori, S., Hamburger, K., & Spillmann, L. (2006). Reversal of apparent rotation in the Enigma-figure with and without motion adaptation and the effect of T-junctions. *Vision Research*, *46*(19), 3267–3273.

Gori, S., Mascheretti, S., Giora, E., Ronconi, L., Ruffino, M., Quadrelli, E., et al. (2015). The DCDC2 intron 2 deletion impairs illusory motion perception unveiling the selective role of magnocellular-dorsal stream in reading (dis)ability. *Cerebral Cortex*, *25*(6), 1685–1695.

Gori, S., Molteni, M., & Facoetti A. (2016). Visual Illusions: An Interesting Tool to Investigate Developmental Dyslexia and Autism Spectrum Disorder. *Frontiers in Human Neuroscience 10:175*.

Gori, S., Pedersini, R., & Giora, E. (2008). How do painters represent motion in garments? Graphic invariants across centuries. *Spatial Vision*, *21*(3–5), 201–227.

Gori, S., Seitz, A.R., Ronconi, L., Franceschini, S., & Facoetti, A. (2016). Multiple causal links between magnocellular-dorsal pathway deficit and developmental dyslexia. *Cerebral Cortex*, *26*(11), 4356–4369.

Gori, S., & Stubbs, D.A. (2014) Motion illusions as a psychophysical tool to investigate the visual system. In Geremek A., Greenlee M.W., Magnussen S. (Eds.), Perception beyond Gestalt: Progress in vision research (pp. 128–143). New York, U.S.A.: Psychology Press.

Gori, S., & Yazdanbakhsh, A. (2008) "The riddle of the Rotating-Tilted-Lines illusion." *Perception*, *37*(4), 631–635.

Grossberg, S., & Mingolla, E. (1993). Neural dynamics of motion perception: Direction fields, apertures, and resonant grouping. *Perception & Psychophysics*, *53*(3), 243–278.

Grossberg, S., Mingolla, E., & Viswanathan, L. (2001). Neural dynamics of motion integration and segmentation within and across apertures. *Vision Research*, *41*(19), 2521–2553.

Guilford, J. P. (1929). Illusory movement from a rotating barber pole. *American Journal of Psychology*, *41*, 686–687.

Gurnsey, R., & Pagé, G. (2006). Effects of local and global factors in the Pinna illusion. *Vision Research*, *46*(11), 1823–1837.

Gurnsey, R., Sally, S. L., Potechin, C., & Mancini, S. (2002). Optimising the Pinna--Brelstaff illusion. *Perception*, *31*(10), 1275–1280.

Hamburger, K. (2007). Apparent rotation and jazzing in Leviant's Enigma illusion. *Perception*, *36*(6), 797–807.

Hamburger K. (2012). Still motion? Motion illusions and luminance contrast. *Perception*, *41*, 113–116.

Hine, T. J., Cook, M., & Rogers, G. T. (1995). An illusion of relative motion dependent upon spatial frequency and orientation. *Vision Research*, *35*(22), 3093–3102.

Hine, T. J., Cook, M., & Rogers, G. T. (1997). The Ouchi illusion: An anomaly in the perception of rigid motion for limited spatial frequencies and angles. *Perception & Psychophysics*, *59*(3), 448–455.

Ito, H. (2012). Illusory object motion in the centre of a radial pattern: the pursuit–pursuing illusion. *i-Perception*, *3*(1), 59–87.

Leviant, I. (1996). Does "brain-power" make Enigma spin? *Proceedings of the Royal Society of London, B*, *263*(1373), 997–1001.

Lidén, L., & Mingolla, E. (1998). Monocular occlusion cues alter the influence of terminator motion in the barber pole phenomenon. *Vision Research*, *38*(24), 3883–3898.

Lorenceau, J., Shiffrar, M., Wells, N., & Castet, E. (1993). Different motion sensitive units are involved in recovering the direction of moving lines. *Vision Research*, *33*(9), 1207–1217.

Kim, S. H., & Papathomas, T. V. (2011). Nonrigid illusory motion in depth induced by translational motion of static images. *i-Perception*, *2*(2), 154–158.

Kitaoka, A., & Ashida, H. (2007). A variant of the anomalous motion illusion based upon contrast and visual latency. *Perception*, *36*(7), 1019–1035.

Kumar, T., & Glaser, D. A. (2006). Illusory motion in Enigma: a psychophysical investigation. *Proceedings of the National Academy of Sciences USA*, *103*(6), 1947–1952.

Kuriki, I., Ashida, H., Murakami, I., & Kitaoka, A. (2008). Functional brain imaging of the rotating snakes illusion by fMRI. *Journal of Vision*, *8*(10):16, 1–10.

Mackay, D. M. (1957). Some further visual phenomena associated with regular patterned stimulation. *Nature*, *180*(4595), 1145–1146.

Mather, G. (2000). Integration biases in the Ouchi and other visual illusions. *Perception*, *29*(6), 721–727.

Mingolla, E., Todd, J. T., & Norman, J. F. (1992). The perception of globally coherent motion. *Vision Research*, *32*(6), 1015–1031.

Mon-Williams, M., & Wann, J. P. (1996). An illusion that avoids focus. *Proceedings of the Royal Society of London, B*, *263*(1370), 573–578.

Murakami, I., Kitaoka, A., & Ashida, H. (2006). A positive correlation between fixation instability and the strength of illusory motion in a static display. *Vision Research*, *46*(15), 2421–2431.

Naor-Raz, G., & Sekuler, R. (2000). Perceptual dimorphism in visual motion from stationary patterns. *Perception*, *29*(3), 325–335.

Otero-Millan, J., Macknik, S. L., & Martinez-Conde, S. (2012). Microsaccades and blinks trigger illusory rotation in the rotating snakes illusion. *Journal of Neuroscience*, *32*, 6043–6051.

Pack, C. C., & Born, R. T. (2001). Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. *Nature*, *409*(6823), 1040–1042.

Pack, C. C., Livingstone, M. S., Duffy, K. R., & Born, R. T. (2003). End-stopping and the aperture problem: Two-dimensional motion signals in macaque V1. *Neuron*, *39*(4), 671–680.

Pinna, B., & Brelstaff, G. J. (2000). A new visual illusion of relative motion. *Vision Research*, *40*(16), 2091–2096.

Pinna, B., & Spillmann, L. (2005). New illusions of sliding motion in depth. *Perception*, *34*(12), 1441–1458.

Ruzzoli, M., Gori, S., Pavan, A., Pirulli, C., Marzi, C. A., & Miniussi, C. (2011). The neural basis of the Enigma illusion: A transcranial magnetic stimulation study. *Neuropsychologia*, *49*(13), 3648–3655.

Spillmann, L. (2009). Phenomenology and neurophysiological correlations: Two approaches to perception research. *Vision Research*, *49*(12), 1507–1521.

Spillmann, L. (2013). "The Ouchi-Spillmann illusion revisited." *Perception*, *42*(4), 413–429.

Spillmann, L., Heitger, F., & Schuller, S. (1986). Apparent displacement and phase unlocking in checker-board patterns. In *Proceedings of the 9th European Conference on Visual Perception*. Bad Nauheim, Germany.

Spillmann, L., & Werner, J. S. (1990). *Visual perception: The neurophysiological foundations*. San Diego, CA: Academic Press.

Troncoso, X. G., Macknik, S. L., Otero-Millan, J., & Martinez-Conde, S. (2008). Microsaccades drive illusory motion in the Enigma illusion. *Proceedings of the National Academy of Sciences USA*, *105*(41), 16033–16038.

Wallach, H. (1935). Über visuell wahrgenommene Bewegungsrichtung. *Psychologische Forschung*, *20*, 325–380 (English translation: S. Würger, R. Shapley, & N. Rubin [1996]. On the visually perceived direction of motion by Hans Wallach: 60 years later. *Perception*, *25*(11), 1317–1367)

Yazdanbakhsh, A., & Gori, S. (2008). A new psychophysical estimation of the receptive field size. *Neuroscience Letters*, *438*(2), 246–251.

Yazdanbakhsh, A., & Gori, S. (2011). Mathematical analysis of the accordion grating illusion: A differential geometry approach to introduce the 3D aperture problem. *Neural Networks*, *24*(10), 1093–1101.

Zeki, S., Watson, J. D., & Frackowiak, R. S. (1993). Going beyond the information given: The relation of illusory visual motion to brain activity. *Proceedings of the Royal Society of London, B*, *252*(1335), 215–222.

Chapter 67
The Enigmatic Enigma Illusion

Kai Hamburger

THE EFFECT

The Enigma illusion is a pattern that in its original version consists of 120 black radial lines on a white background intercepted by three bicolored annuli and a central disk. The main illusory effect in the Enigma (leading to its name) occurs during fixation of the center of the static image. Then quite intense streaming motion may be perceived on the different annuli. It is characterized by a traveling wave or some subtle motion on the annuli that may not be described in more detail by the observer. Sometimes the observers call it "a feeling of motion." This perceived (illusory) motion can occur in either direction; clockwise or counterclockwise. From time to time the direction of the streaming motion is reversed (Gori, Hamburger, & Spillmann, 2006).

This illusion was first introduced by Isia Leviant in 1981 (Leviant, 1996; a variation is shown in Figure IV.67-1). Here a version with colors ranging from dark blue to light blue is shown, while the pattern may also consist of "rainbow" colors and many other color combinations (e.g., Fig. IV.67-3). However, the different colors and color combinations contribute only to the aesthetic appeal of this configuration but not to the genuine illusory effect, which is demonstrated later. Inspection of the Enigma pattern might/should also lead to a percept of shimmering or flickering of the high spatial frequency radial lines (MacKay, 1957a, 1957b). This, however, is just a side effect. Additionally, many observers report that the classical Enigma pattern by Leviant and some of its variants produce something akin to a "tube" (depth perception). Other perceptual effects may occur upon observation of the Enigma figure or variants thereof, but those described here constitute the main effects that are reliably connected to this stimulus.

RELEVANCE

This illusion is one of three other static motion illusion patterns that have been investigated (the Pinna-Brelstaff, rotating snakes, and rotating-tilted-lines illusions; see chapters 74 and 66 in this volume) for whether its perception is affected by age (Billino, Hamburger, & Gegenfurtner, 2009). It has been found that even young children at the age of three years perceive the streaming motion on the three bipartite annuli (more than 75% of the participants). For adults there is a slight increase in the perception of this effect (~80%), which is not affected by aging later on. This illusion as well as the Pinna-Brelstaff illusion, in contrast to the rotating snakes and the rotating-tilted-lines illusion, remain stable throughout life, even though about one-quarter of the population does not perceive—or, rather, is not tricked by—this illusion.

Such illusions are an important means to psychophysically investigate human motion perception and its limits. The illusions are currently investigated with brain-imaging techniques to further explore how the different aspects (motion, color, luminance) are processed (in isolation as well as together) in the human brain.

VARIANTS AND PARAMETERS

A few variants of this static motion illusion pattern have been created in recent years. For example, an achromatic version is shown in Figure IV.67-2. Even though the Enigma illusion elicits a strong motion percept (in most observers), research on this pattern is to date rather rare.

One of the major and most impressive variants has been introduced by Gori and Hamburger (2006). They created a combination of the Enigma illusion, the MacKay rays, and the Pinna-Brelstaff illusion. The MacKay ray figure consists of high spatial frequency radial lines that seem to shimmer during observation (MacKay, 1957a, 1957b). The Pinna-Brelstaff illusion (Pinna & Brelstaff, 2000) consists of two rings made from multiple diamond shapes (two white and black edges, grey background). During self-motion of the head or body (looming), the two rings perceptually rotate in opposite directions.

The combination of these three illusion patterns also consists of black radial lines, but the lines making up an Enigma-like annulus are tilted in the opposite direction compared to the surrounding area (central area and outer surround). This illusion is now known as the rotating-tilted-lines illusion (Gori & Hamburger, 2006). A variation of this illusion, which was discovered before the rotating-tilted-lines illusion was introduced, includes the three bipartite annuli of the Enigma, while the radial lines are tilted about 45° clockwise and counterclockwise in alternation in the areas between the colored annuli (Figures IV.67-3 and IV.67-4, left).

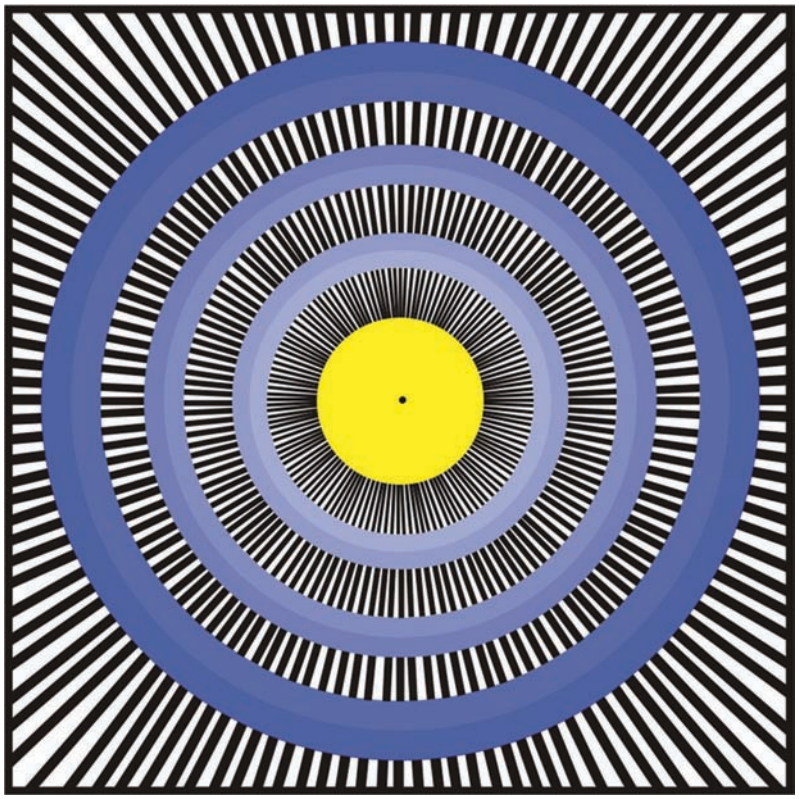


Figure IV.67-1. A variant of the original Enigma illusion (reproduction by K. Hamburger). The Enigma pattern consists of high-frequency black radial lines with three bipartite-colored annuli and a central disk superimposed onto the radial lines. The radial lines seem to shimmer, while streaming motion in clockwise and/or counterclockwise direction may be perceived on the colored annuli.

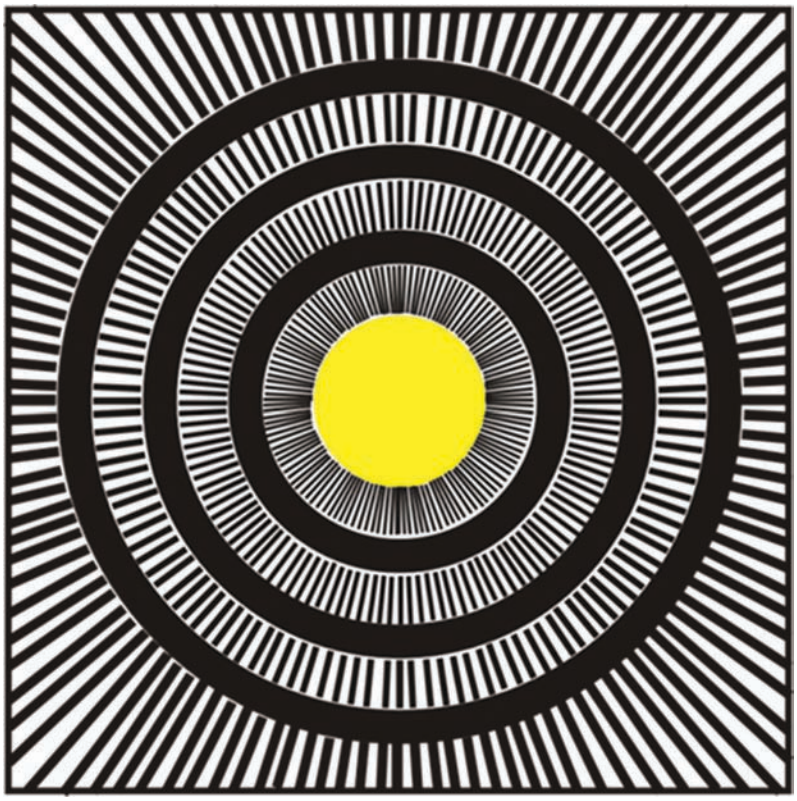


Figure IV.67-2. This figure shows that color is not fundamental for the streaming motion in the Enigma illusion, since this black/white variant also leads to the characteristic motion percept (the yellow center may also be replaced by a black disk without affecting the streaming motion). Some "light" (gray) areas seem to move on the rings in either direction. It may be concluded that color per se enhances the aesthetic appeal of these illusions but is not a prerequisite for the illusory motion.

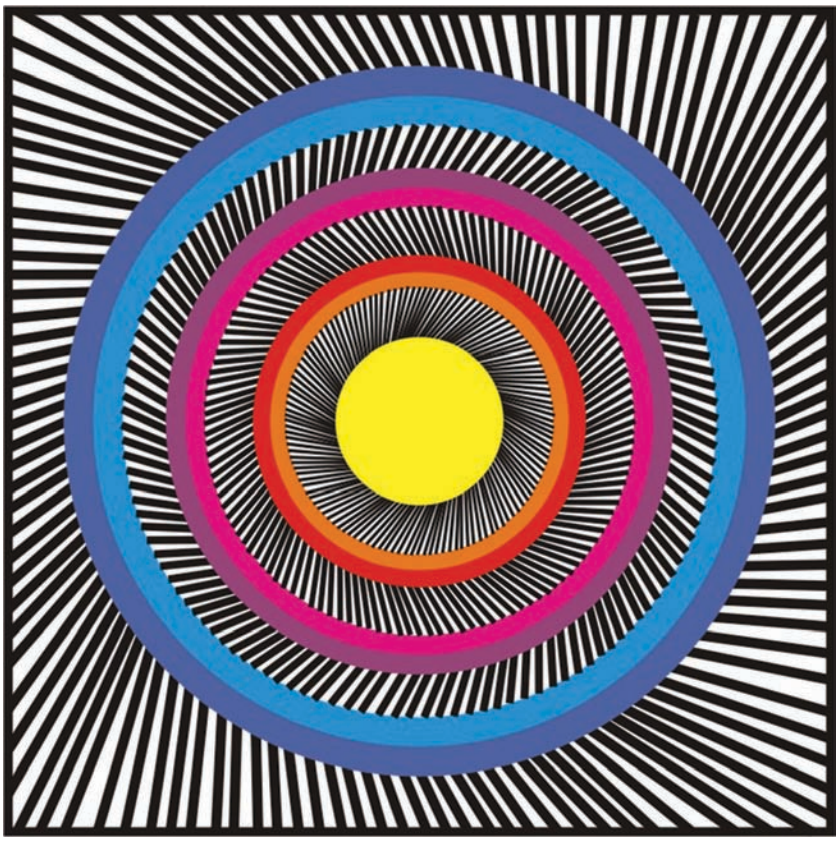


Figure IV.67-3. This pattern represents a combination of two illusions, namely the Enigma and the Pinna-Brelstaff illusion (Pinna & Brelstaff, 2000). Streaming motion as described before may be perceived on the differently colored annuli while the tilted high spatial frequency black lines seem to rotate in opposite directions during body and/or head movements back and forth (toward and away from the stimulus pattern; looming). This second illusory effect is a variation of the rotating-tilted-lines illusion by Gori and Hamburger (2006).

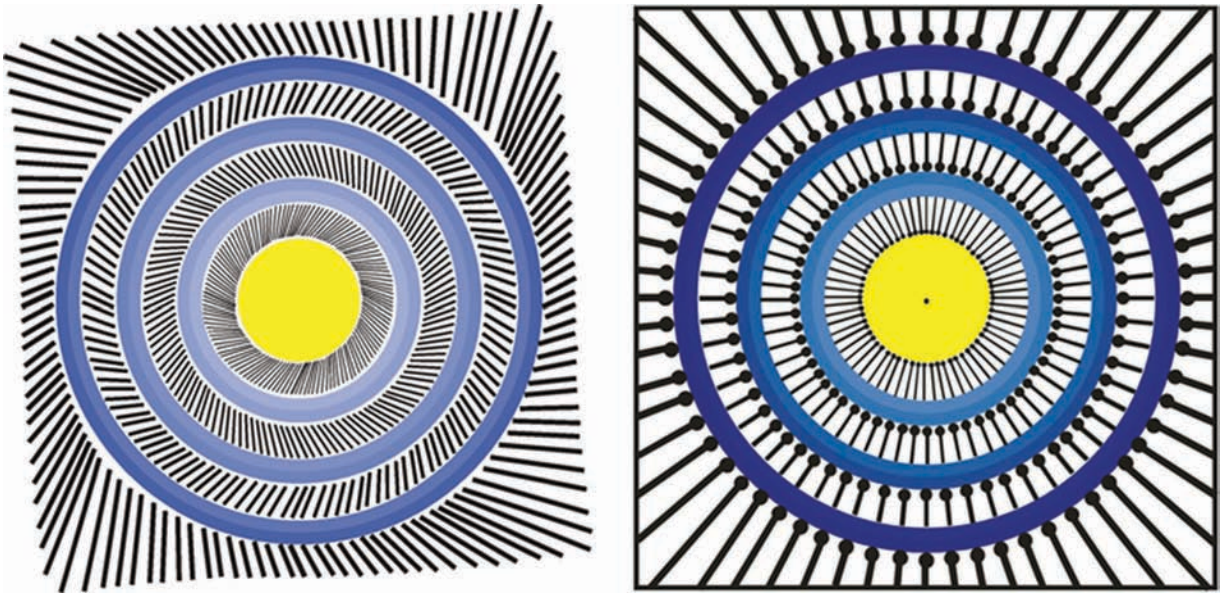


Figure IV.67-4. Left: Another appealing variant of the Enigma illusion. On the one hand, due to the tilt and the nonadjacent lines, it represents a quite aesthetic variant. From a scientific point of view it is valuable, since the streaming motion still occurs, even though the high spatial frequency lines do not touch the rings (the T-junction hypothesis for explaining the streaming motion; e.g., Gori et al., 2006). Right: A further Enigma variant for the investigation of the influence of T-junctions (orientation-selective neurons in the brain) for the occurrence of streaming motion. A black blob was inserted at each position where a part of a radial line is adjacent to the ring (inward direction). This modification breaks up the T-junctions in the Enigma figure; a T-junction is normally present where a radial line touches a ring, creating a "T" (the leg of the T is represented by the radial line, the top is represented by the ring). The streaming motion on the rings is attenuated here.

For the rotating-tilted-lines illusion as well as for this Enigma variant, a perceived motion of the radial lines is encountered when the observer moves toward and away from the pattern. In contrast to the perceived motion on the classical Enigma annuli—which is slightly attenuated in the two variants shown here—the motion direction (clockwise or counterclockwise) of the tilted high spatial frequency lines depends on whether the pattern becomes smaller or larger in the visual field of the observer (due to body or head movements toward and away from the pattern).

For the original Enigma illusion, some other parameters and variants have been tested empirically. Kumar and Glaser (2006) examined the optimal conditions for the streaming motion to occur, while Hamburger (2007) tested the limits of this illusion, namely the conditions under which the illusion breaks down. Thus far it can be said that the strongest percept of streaming motion occurs in versions that come close to the original version by Leviant (1996). However, changing the spatial frequency of the radial lines or replacing them with other elements significantly reduces the effects but does not totally abolish them. For example, Hamburger (2012) tested the effect of luminance contrast and equiluminance on the Enigma and three other static motion illusions (Pinna-Brelstaff, rotating snakes, rotating-tilted-lines). The illusion was strongest at high luminance contrast (black/white as in the original), slightly reduced at low luminance contrast (gray lines and annuli in front of a slightly lighter gray background), and even more reduced under conditions of equiluminance (different hues but same lightness). However, a residual streaming motion persists under conditions of equiluminance. Such findings are in contrast to empirical data obtained for the so-called geometric-optical illusions (see chapter 18 in this volume). This class of visual illusions remains stable under equal luminance conditions (Hamburger, Hansen, & Gegenfurtner, 2007).

DISCUSSION AND CONCLUSIONS

Many different explanations for this impressive illusory effect have been put forward recently, ranging from retinal explanations (e.g., accommodation, microsaccades, saccadic eye movements) to cortical explanations (e.g., end-stopped neurons at T-junctions; motion sensitive areasV5/V5+ and MST). Related variants of the Enigma concerning the T-junction hypothesis (e.g., Gori et al., 2006) are demonstrated in Figure IV.67-4. For a more detailed overview of possible underlying mechanisms, see, for example, Hamburger (2007) and Ruzzoli et al. (2011).

Taken together, the Enigma constitutes a pattern that requires different factors to occur. At this point more research is required to unveil the underlying mechanisms of the still enigmatic Enigma.

REFERENCES

Billino, J., Hamburger, K., & Gegenfurtner, K. R. (2009). Age effects on the perception of motion illusions. *Perception*, *38*, 508–521.

Gori, S., & Hamburger, K. (2006). A new motion illusion: The rotating-tilted-lines illusion. *Perception*, *35*, 853–857.

Gori, S., & Hamburger, K., & Spillmann, L. (2006). Reversal of apparent rotation in the Enigma-figure with and without motion adaptation and the effect of T-junctions. *Vision Research*, *46*, 3267–3273.

Hamburger, K. (2007). Apparent rotation and jazzing in Leviant's Enigma illusion. *Perception*, *36*, 797–807.

Hamburger, K. (2012). Still motion? Motion illusions and luminance contrast. *Perception*, *41*(1), 113–116.

Hamburger, K., Hansen, T., & Gegenfurtner, K. R. (2007). Geometric-optical illusions at isoluminance. *Vision Research*, *47*, 3276–3285.

Kumar, T., & Glaser, D. A. (2006). Illusory motion in Enigma: A psychophysical investigation. *Proceedings of the National Academy of Sciences USA*, *103*, 1947–1952.

Leviant, I. (1996). Does "brain-power" make Enigma spin? *Proceedings of the Royal Society of London, B*, *263*, 997–1001.

MacKay, D. M. (1957a). Moving visual images produced by regular stationary patterns. *Nature*, *180*, 849–850.

MacKay, D. M. (1957b). Some further visual phenomena associated with regular patterned stimulation. *Nature*, *180*, 1145–1146.

Pinna, B., & Brelstaff, G. J. (2000). A new visual illusion of relative motion. *Vision Research*, *40*, 2091–2096.

Ruzzoli, M., Gori, S., Pavan, A., Pirulli, C., Marzi, C. A., & Miniussi, C. (2011). The neural basis of the Enigma illusion: A transcranial magnetic stimulation study. *Neuropsychologia*, *49*, 3648–3655.

Chapter 68
The Fraser-Wilcox Illusion and Its Extension

Akiyoshi Kitaoka

INTRODUCTION

This chapter reviews the Fraser-Wilcox illusion, an anomalous motion illusion displayed as a stationary image, and its extension, for which I have produced a variety of illusion works (see Kitaoka, 2015). The Internet has made one of these (Fig. IV.68-1) popular with the public. This illusion is sometimes called the rotating snakes illusion (e.g., Billino, Hamburger, & Gegenfurtner, 2009; Kuriki, Ashida, Murakami, & Kitaoka, 2008), but I suggest that this illusion should be an "optimized version" of the Fraser-Wilcox illusion rather than a new illusion.

ANOMALOUS MOTION ILLUSIONS

A variety of visual illusions involving shape, color, lightness, depth, motion, faces, and so on have been discovered since the 19th century. In the history of the study of motion illusion, most stimuli, needless to say, have been moving

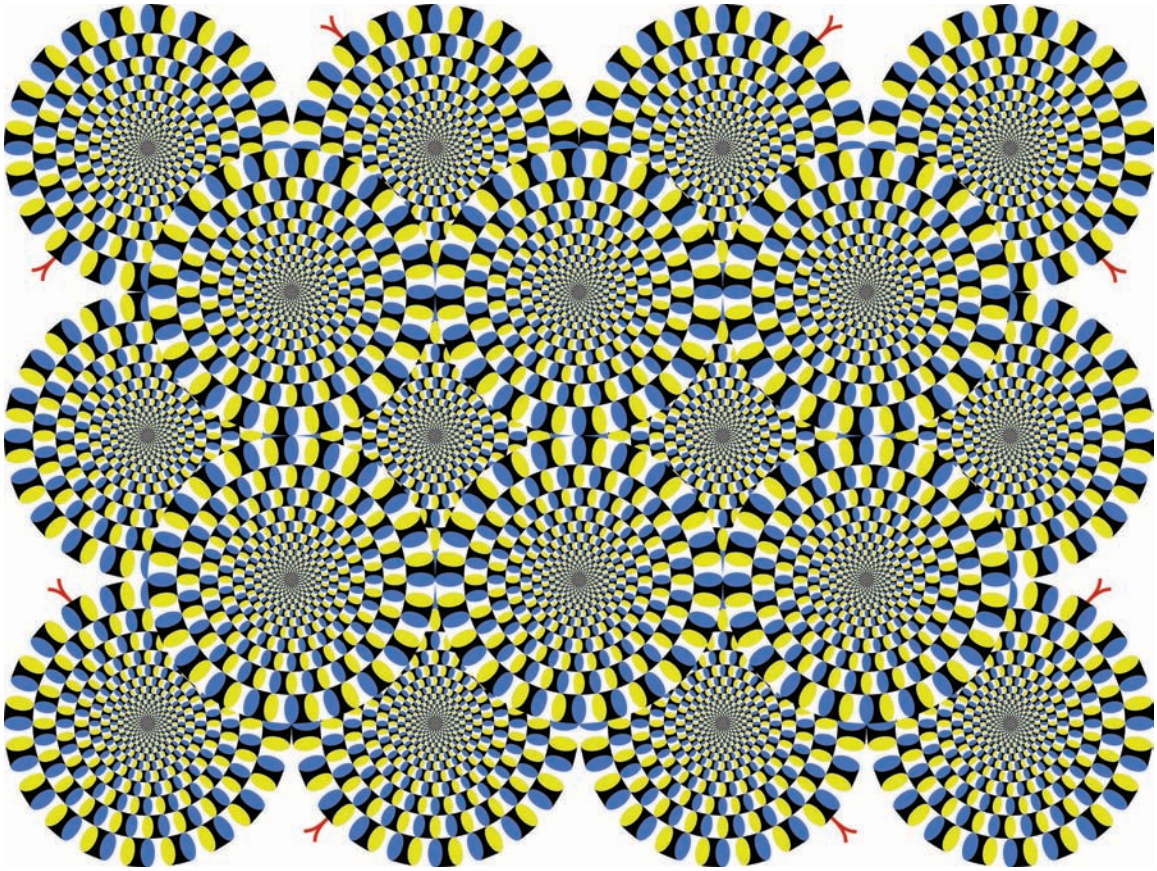


Figure IV.68-1. The rotating snakes illusion, which I produced in 2003 and have exhibited on the Internet. Each disk appears to rotate automatically in the constant direction (i.e., from black, blue, white, yellow, and back to black). (Figure by Akiyoshi Kitaoka.)

objects, for example a fall in the waterfall illusion (motion aftereffect; Mather, Verstraten, & Anstis, 1998), moving rectangles in the footstep illusion (Anstis, 2000), or moving ellipses in the swimmers illusion (Shapiro, Knight, Palmor, & Mancusi, 2006). On the other hand, the study of anomalous motion illusions, which feature the use of stationary stimuli, started in the middle of the 20th century and has accelerated since the 1980s.

Three classes of anomalous motion illusions can be distinguished (Kitaoka & Ashida, 2007). The first class refers to *scintillating* phenomena, which are often observed in op-art, for example the MacKay (1957) ray pattern, the Enigma painting (Leviant, 1996), Bridget Riley's "Fall" (Zanker, Doyle, & Walker, 2003), and others. The second class is characterized by *sliding* phenomena, which are induced by the retinal slip of an image. This class consists of two subclasses. One is characterized by an illusory motion, the direction of which is different from or perpendicular to the retinal slip. This subclass includes the Ouchi-Spillmann illusion (Spillmann, 2013; Spillmann, Heitger, & Schuller, 1986), the Hine illusion (Hine, Cook, & Rogers, 1995), the Pinna illusion (Pinna & Brelstaff, 2000), and the drifting spines illusion (Kitaoka, 2010b), among others. The other is characterized by another illusory motion, the direction of which is the same as or opposed to the direction of the retinal slip. This subclass includes the fluttering-heart illusion (Helmholtz, 1867/1962), the floating-motion illusion (Pinna & Spillmann, 2002), the visual-delay-dependent motion illusion (Kitaoka & Ashida, 2007), and others. The third class refers to *automatically moving* phenomena, in which illusory motion appears in a constant direction guided by patterns. This class includes the Fraser-Wilcox illusion (Fraser & Wilcox, 1979) and its extension reviewed in this chapter, the central drift illusion (Fermüller, Ji, & Kitaoka, 2010; Kitaoka & Ashida, 2004), the drifting arrows illusion (Kitaoka, 2010a), and others.

THE FRASER-WILCOX ILLUSION

Fraser and Wilcox (1979) were the first to propose the third class of anomalous motion illusion. They called it the "escalator illusion" because of the cochlea-like appearance of the image (Fig. IV.68-2). This pattern is made up of pie slices filled with repeated luminance gradients and appears to rotate "automatically" when observers see the image. This image was painted by Alex Fraser (1923–2002), who was a geneticist known as "one of the first to conceive and execute computer simulations of genetic systems, and [his] efforts in the 1950s and 1960s had a profound impact on computational models of evolutionary systems" (Fogel, 2002, p. 429). He was a painter too. Some of his paintings can be seen at http://doctoralexfraser.blogspot.jp/. According to this site, the escalator painting was on canvas and was the one he used for testing the motion effect.

Fraser and Wilcox (1979) reported a perceptual dimorphism—that is, some observers saw motion in the direction of light shading while others reported motion in the direction of dark shading. A parent–child survey suggested a genetic origin for the observer's directional perception. However, several studies that examined the Fraser-Wilcox illusion failed to reproduce perceptual dimorphism and concluded that only the illusory motion from dark to light shading exists (Cloete, Wallis, Walterso, & Broerseo, 2004; Faubert & Herbert, 1999; Naor-Raz & Sekuler, 2000). Strictly speaking, Faubert and Herbert (1999) reported one exceptional observer from more than 200 observers; this person saw the illusory motion from light to dark shading. Naor-Raz and Sekuler (2000) stated that they had confirmed perceptual dimorphism, yet they only reported the illusion from dark to light shading.

The Fraser-Wilcox illusion is stronger in the peripheral vision than in the central vision (Faubert & Herbert, 1999; Fraser & Wilcox 1979; Naor-Raz & Sekuler, 2000). Faubert and Herbert claimed that this illusion is triggered by eye movement or a blink and called it the peripheral drift illusion. They assumed that this illusory motion is based on the difference in visual latency between light and dark parts in luminance gradients. Since light parts would give a shorter latency than dark ones, light–dark borders should have a larger timing difference than luminance gradients, thus producing a stronger motion signal across the border. Naor-Raz and Sekuler revealed that the illusion magnitude is a positive, nearly linear function of contrast. Moreover, testing an aphake, they ruled out fluctuations of accommodation as a possible source, as did Cloete et al. (2004).

DISAMBIGUATION OF THE FRASER-WILCOX ILLUSION BY DARKENING OR LOWERING CONTRAST

I recently found that the illusory rotation from dark to light shading is made clear by darkening the whole image (Kitaoka, 2012). Compare the basic image (Fig. IV.68-3a) with its darkened one (Fig. IV.68-3b). The illusion from light to dark shading is hidden or suppressed. This effect is strong when the background is of an intermediate luminance (gray) of the shading. I suggest that the luminance arrangement of three areas meeting at a T-junction, not luminance gradients, is the critical component. When the two areas flanking the stem are light and dark, respectively, and the root side is of an intermediate luminance, illusory motion is directed from light to dark. This idea is embodied in Fig. IV.68-3c, which does not comprise luminance gradients but blocks of three luminance areas.

Darkening an image is inevitably accompanied by lowered contrast. It seems that lowering contrast alone can serve to emphasize the illusory rotation from dark to light shading (Fig. IV.68-3d). This observation disagrees with Naor-Raz and Sekuler's (2000) finding that the illusion magnitude is a positive, nearly linear function of stimulus contrast. This contradiction could be solved by assuming the illusions shown in Figures IV.68-3b, IV.68-3d, and IV.68-3e are different from that of Naor-Raz and Sekuler (2000), who used a light background.

Naor-Raz and Sekuler (2000) used a color image for a screening test prior to the formal test, in which the disk was dark and surrounded by a gray area (Fig. IV.68-3e). This condition agrees with the condition for the darkening disambiguation discussed here. Moreover, Cloete et al. (2004) adopted an image surrounded by a gray background, and the test stimulus was relatively dark (mean: 26 cd/m^2)

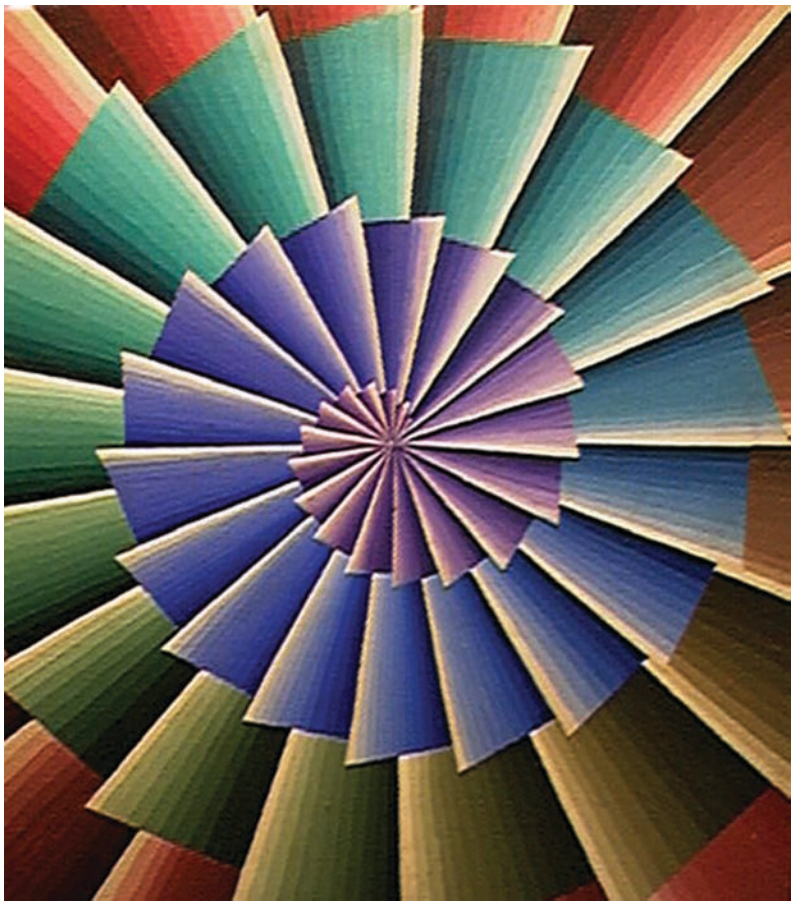


Figure IV.68-2. Alex Fraser's original painting (Fraser, 2006). According to Fraser and Wilcox (1979), it is expected that some observers see the disk rotate clockwise while others see the opposite. (Reproduced with permission of Alan Fraser.)

according to the authors' unpublished handout. This condition also accords with the darkening disambiguation.

PERCEPTUAL DIMORPHISM

The Fraser-Wilcox illusion appears differently depending on whether the background luminance is light or dark. The illusory rotation from dark to light shading is frequently observed when the luminance of the background is the same or higher than the highest luminance in the gradients (Figure IV.68-4a), whereas the illusory rotation from light to dark shading is predominant when the luminance of the background is the same or lower than the lowest luminance in the gradients (Fig. IV.68-4b). My suggestion is that these two opposing illusions might have contributed to perceptual dimorphism or have cancelled each other out.

Thus the Fraser-Wilcox illusion includes at least three different illusions, two being illusions from dark to light shading (Figures IV.68-3b and IV.68-4a) and the other an illusion from light to dark shading (Figure IV.68-4b). This suggestion would also be supported by the finding of "substantial individual variation in the perception of Fraser-Wilcox stimuli, but little variation in the snakes illusion" (Becker & Mollon, 2007, p. 62).

In addition, perceptual dimorphism is observed in particular color combinations. Dark to light shading from purple to red via red-purple surrounded by a red-purple background gives an illusory motion in the direction from dark to light shading (Figure IV.68-5a). On the other hand, light to dark shading from light red-purple (magenta) to red-purple via red surrounded by a red background renders an illusory motion in the direction from light to dark shading (Figure IV.68-5b). Moreover, gradients may not be indispensable for these illusions (see Figs. IV.68-5c and IV.68-5d) like darkening disambiguation. This issue is discussed again later in the chapter.

THE OPTIMIZED FRASER-WILCOX ILLUSION

Kitaoka and Ashida (2003) proposed a new painting rule to draw images that can demonstrate a much stronger illusion (Figure IV.68-6a) than previous Fraser-Wilcox patterns. The proposed rule was a four-stroke luminance combination, and the order is black, dark gray, white, light gray, and back to black. When this combination is repeated, a strong illusion can be observed in this direction. The rotating snakes (Figure IV.68-1) illusion simply follows this rule.

Figure IV.68-6b shows its shading variation, which Kitaoka (2006b, 2007) called the "optimized Fraser-Wilcox illusion, Type I." The illusion shown in this figure is a composite between the illusion from dark to light shading with a lighter surround (Figure IV.68-6c) and the illusion from light to dark shading with a darker surround (Fig. IV.68-6d). These images correspond to Figures IV.68-4a and IV.68-4b, respectively. The limited cases are listed in the top row of

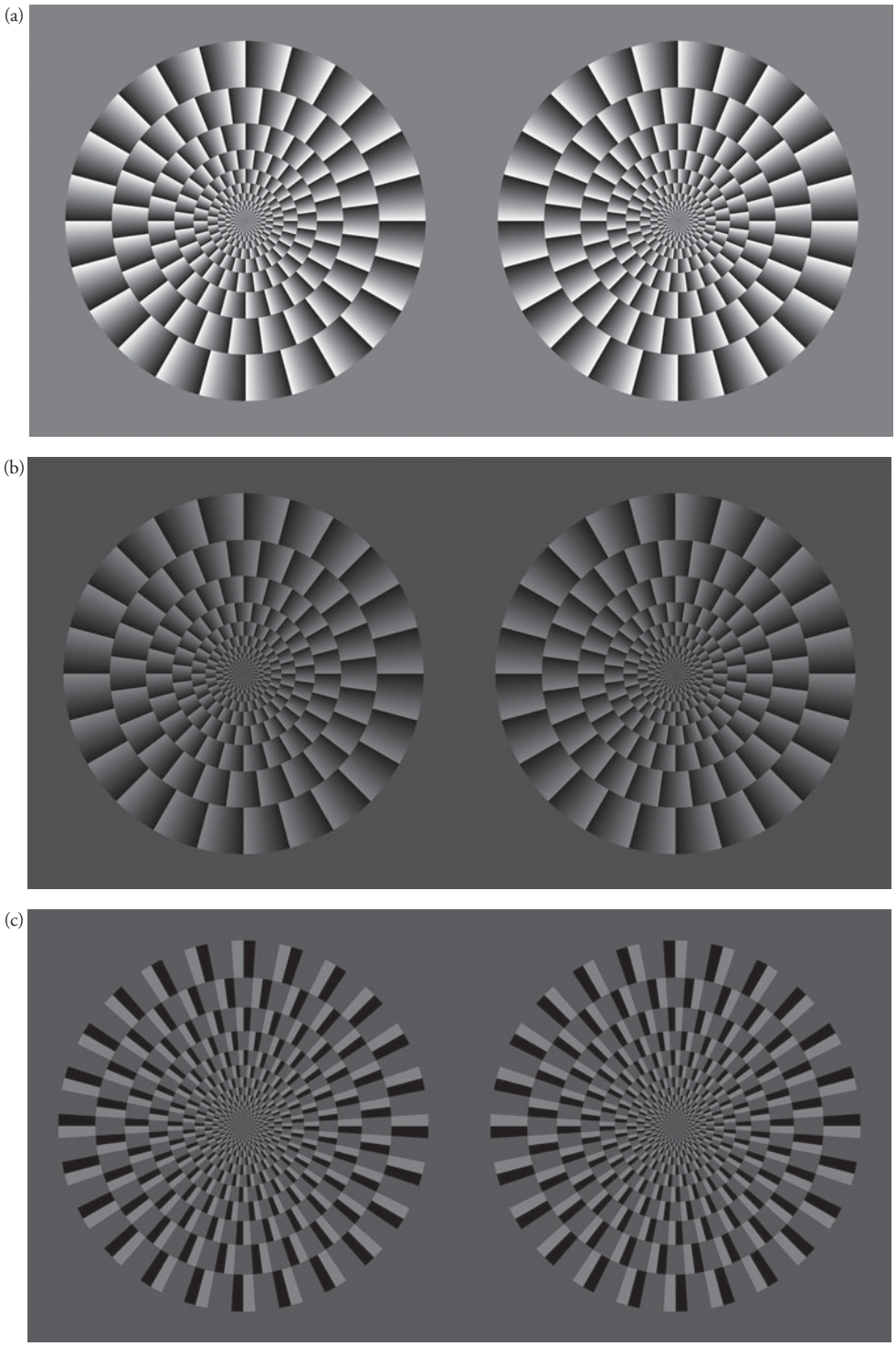


Figure IV.68-3. Disambiguation by darkening or lowering contrast. (a) A simplified, basic image for the Fraser-Wilcox illusion. The direction of perceived rotation appears to be unstable. In the following panels (b) through (e), the left disk appears to rotate clockwise while the right one counterclockwise. (b) Darkening disambiguation of the Fraser-Wilcox illusion by darkening the image (a). (c) A stepwise version of (b), which suggests that luminance gradients are not indispensable for this illusion. (d) Lowering contrast disambiguation of the Fraser-Wilcox illusion through modification of the image (a). (e) Naor-Raz and Sekuler's (2000) screening test image, reproduced with permission of Robert Sekuler. The left half is the original and the right one is its mirror image. Their image could be included in this category. (Figure by Akiyoshi Kitaoka except (e).)

Figure IV.68-7. The luminance of the surround relative to the luminance gradients is critical.

Figure IV.68-6e shows a three-luminance-step illusion, which Kitaoka (2007) called "Type IIa." The basic image consists of three areas of different luminances, with one thinner area flanked by thicker areas (see the Type IIa row in Fig. IV.68-7). The thinner area should be the darkest or the lightest among the three areas. This type also comprises two opposing illusions. One is the illusory motion directed from light gray to dark gray via black; the other

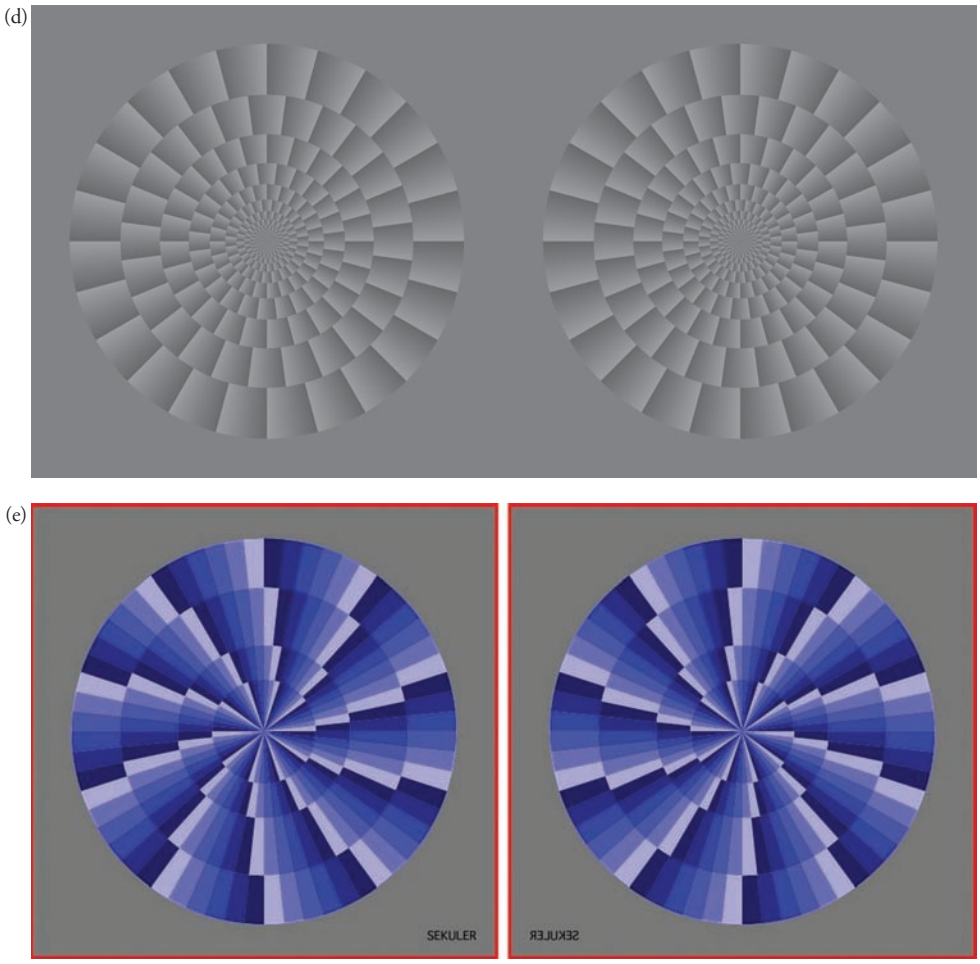


Figure IV.68-3. Continued

is the direction from dark gray to light gray via white. In this connection, Figure IV.68-6f shows another three-luminance-step illusion, which Kitaoka (2007) called "Type IIb," in which the thinner area is of an intermediate luminance between the flankers. When the thinner area is dark gray, the illusory motion is in the direction from black to white via dark-gray; when the thinner area is light gray, the illusory motion is in the direction from white to black via light gray (see the Type IIb row in Fig. IV.68-7). In addition, there is a close, pictorial resemblance in the basic images between Type IIa and "reversed phi movement" (Anstis, 1970; Anstis & Rogers, 1975) and between Type IIb and "phi movement" (Gregory & Heard, 1983; for a review see Kitaoka, 2006a).

Two luminance steps in a row also produce an illusion. Kitaoka (2007) called it "Type III" (Fig. IV.68-6g). This type comprises two illusions. One is the illusory motion directed from black to dark gray when the image is surrounded by a lighter area; the other is the direction from white to light gray when the image is surrounded by a darker area. These limited cases are shown in the Type III row of Figure IV.68-7.

Figure IV.68-6h shows a variation that makes a convex or concave impression because edges are shaded or highlighted so as to generate shape-from-shading phenomena (Horn, 1975; Ramachandran, 1988). Kitaoka (2008a, 2008b) called it "Type IV." This variation comprises two illusions. One is the illusory motion directed from a darker edge to the adjacent dark area; the other is the direction from a lighter edge to the adjacent light area (see the Type IV row of Fig. IV.68-7).

In sum, I propose 10 basic illusions (five types by two directions) as listed in Figure IV.68-7. Since this classification remains at the pictorial level, it is necessary to reveal the underlying mechanisms in the next step.

ENHANCEMENT BY COLOR

Although the optimized Fraser-Wilcox illusion basically depends on the spatial arrangement of luminance, color enhances the illusion. There are two types of color enhancement. For the first type, any color combination will work. According to my experience, red or blue most strongly enhances the illusory motion from dark to light, whereas yellow or green strongly enhances the illusory motion from light to dark. Figure IV.68-8a shows an example that corresponds to a colored image of Type I, in which the dark to light shading is replaced with a black to magenta shading (magenta is a composite of red and blue) while the light to dark shading is replaced with a yellow to yellowish-green shading. This method of color enhancement has been implemented with success in all types of the optimized Fraser-Wilcox illusion (Kitaoka, 2007, 2008a, 2008b, 2010a).

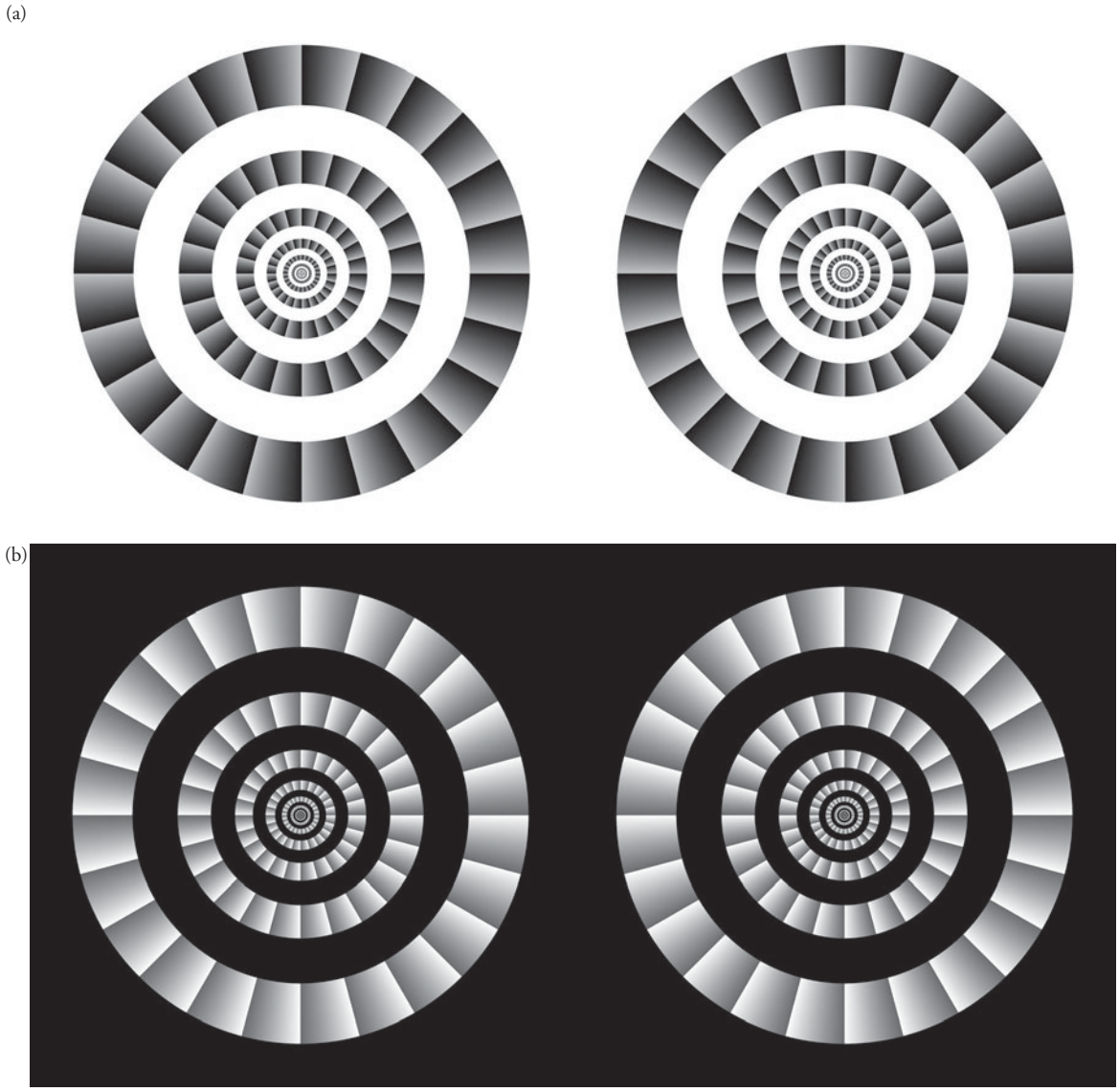


Figure IV.68-4. Background luminance dependency. In each panel, the left disk appears to rotate clockwise while the right one counterclockwise. (a) When the luminance of the background is the same or higher than the highest luminance in the gradients, the direction of illusory motion is from dark to light shading. (b) When the luminance of the background is the same or lower than the lowest luminance in the gradients, the direction of illusory motion is from light to dark shading. (Figure by Akiyoshi Kitaoka.)

For the second type, a combination of red and blue plays a part. Shaul Baskin, chief of Wezit Research Group (http://www.visnsoft.com/), developed a novel, still stronger color enhancement before 2005. He drew my attention to his website by e-mail in August 2005, before which I was not aware of this type of color enhancement. According to my observations, the best combination is the arrangement from purple, red-purple, magenta (light red-purple), red, and back to purple. Purple should be darkest and magenta should be lightest. Figure IV.68-8b shows an example that corresponds to a colored image of Type I. The base color is red and the modulating color is blue, though some of Baskin's works include orange instead of magenta. The color illusions shown in Figure IV.68-5 may also be included in this type.

THE EFFECT OF AGE

Fluctuation of accommodation is unlikely to contribute much to the Fraser-Wilcox illusion. Naor-Raz and Sekuler (2000) examined an 82-year-old bilateral post-op aphake. This person reported the illusion, and his ratings were similar to those of younger participants.

In connection with this observation I collected data from more than 1,000 people to examine whether age is related to the rotating snakes illusion (Fig. IV.68-1). I distributed a high-resolution A4-size print of the picture to participants who came to my lectures or events in schools or museums and asked them to rate the strength of illusion based on the following scale: "I saw no apparent motion" (zero), "I slightly saw apparent motion" (1), "I saw apparent motion" (2), and "I saw strong apparent motion" (3). In total, 1,863 people participated, and their ages were between 2 and 92 years. The results are shown in Table IV.68-1. As a whole, a weak but significant negative correlation between age and illusion was observed ($r = .33$, $df = 1861$, $p < .01$), a result consistent with a previous study (Billino et al., 2009). The result does not support the accommodation hypothesis, because the majority of elder participants, even those in their

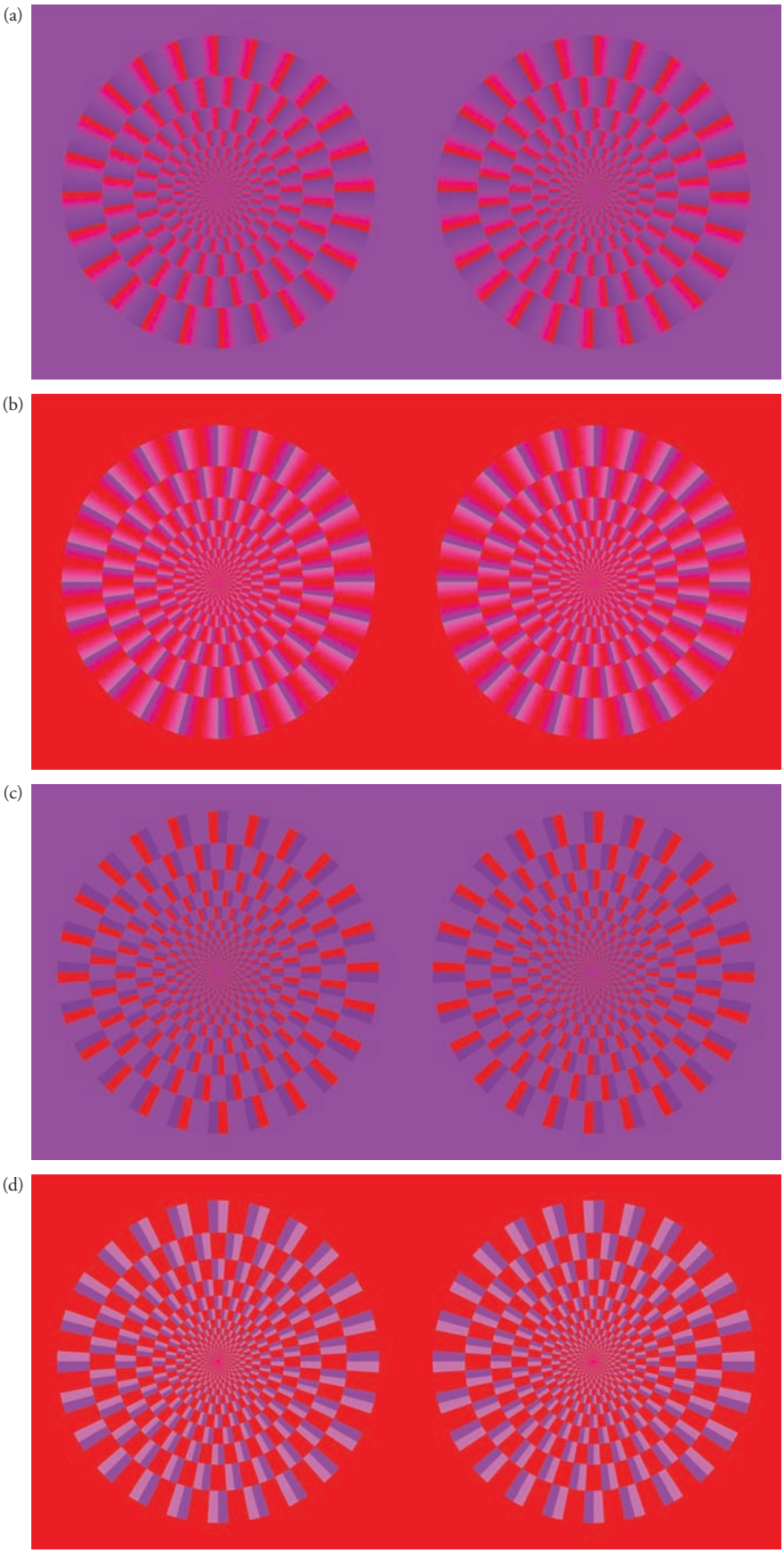


Figure IV.68-5. Color-dependent perceptual dimorphism. In each panel, the left disk appears to rotate clockwise while the right one counterclockwise. (a) A particular color combination promotes the illusion from dark (purple) to light (red) shading. (b) Another particular color combination promotes the illusion from light (light red-purple or magenta) to dark (purple) shading. (c) and (d) Simplified versions of (a) and (b), respectively, which suggest that gradients are not indispensable for these illusions. (Figure by Akiyoshi Kitaoka.)

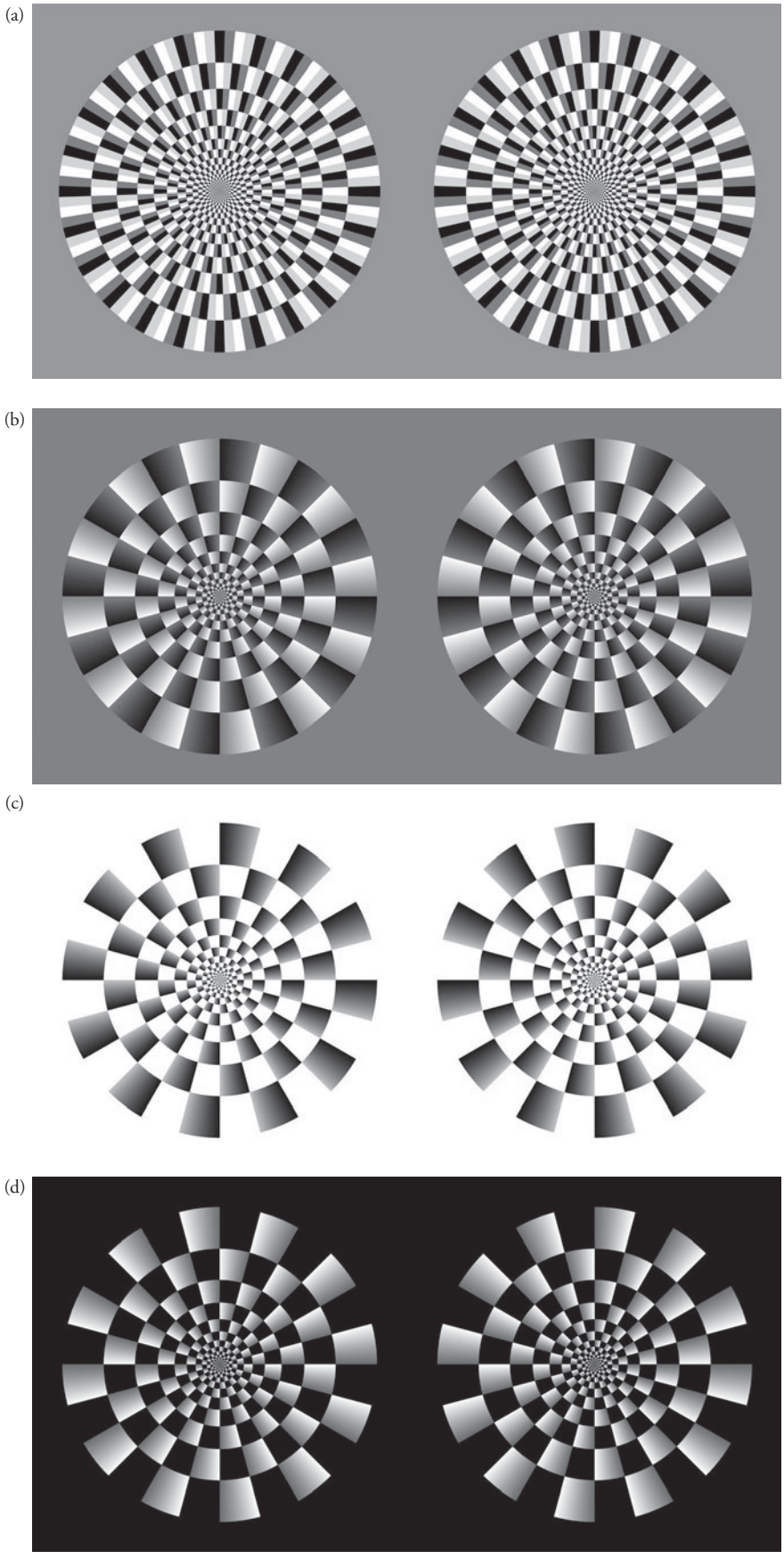
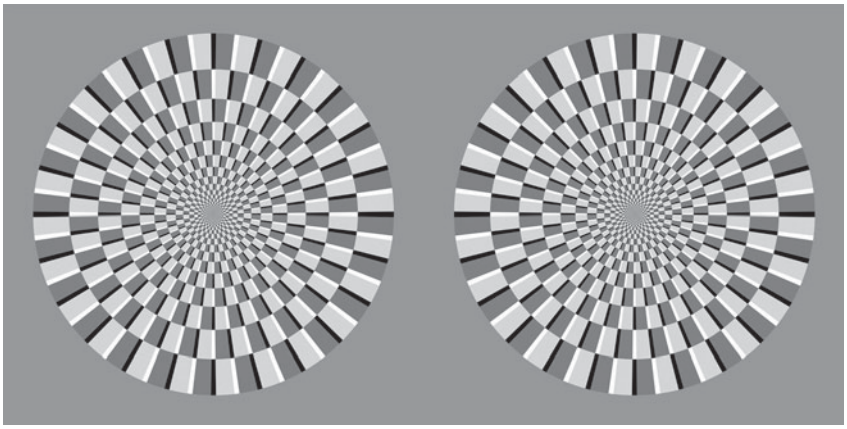
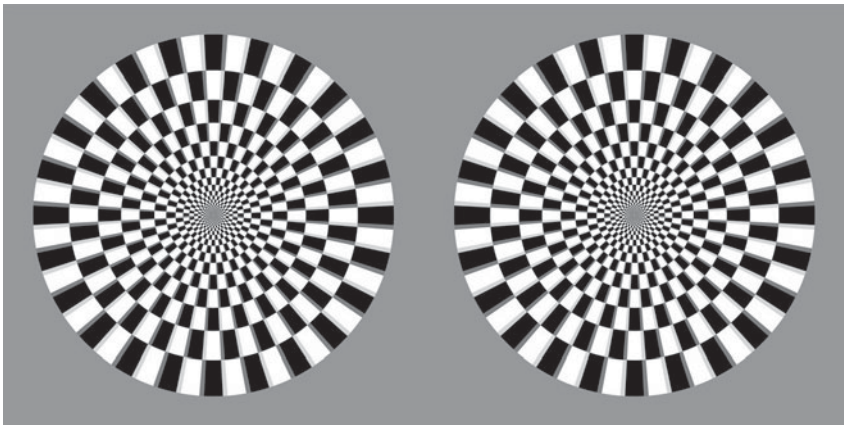


Figure IV.68-6. The "optimized" Fraser-Wilcox illusion. In each panel, the left disk appears to rotate clockwise while the right one counterclockwise. (a) Kitaoka and Ashida's (2003) original version. The illusory motion is directed from black, dark gray, white, light gray, and back to black. (b) Type I, the shading version. The illusory motion is directed from black to gray shading and from white to gray shading. (c) The illusion from dark to light shading within Type I. (d) The illusion from light to dark shading within Type I. (e) Type IIa, in which the basic pattern consists of three areas of different luminance. The illusory motion is in the direction from light gray to dark gray via black or in the direction from dark gray to light gray via white. (f) Type IIb, in which the basic pattern consists of three areas of different luminance. The illusory motion is in the direction from black to white via dark gray or in the direction from white to black via light gray. (g) Type III, in which the basic pattern consists of two areas of different luminance. The illusory motion is in the direction from black to dark gray or in the direction from white to light gray. (h) Type IV, which is observed outdoors at lighted or shaded bulges or concaves. The illusory motion is in the direction from a darker edge to the adjacent darker area or in the direction from a lighter edge to the adjacent lighter area. See Figure IV.68-7. (Figure by Akiyoshi Kitaoka.)

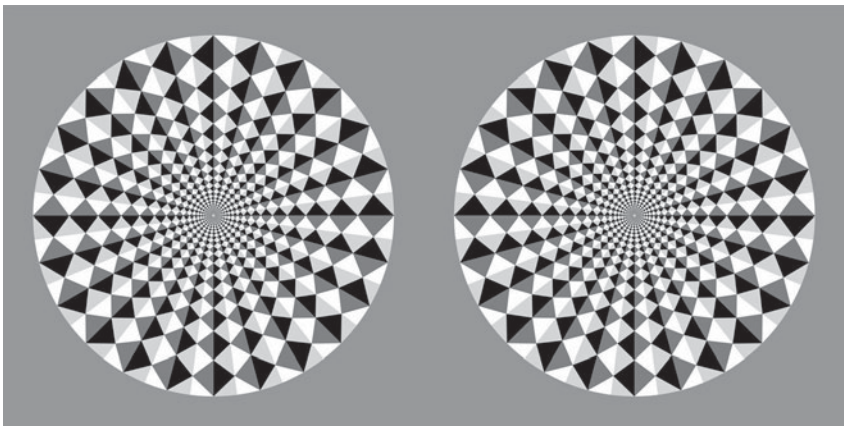
(e)



(f)



(g)



(h)

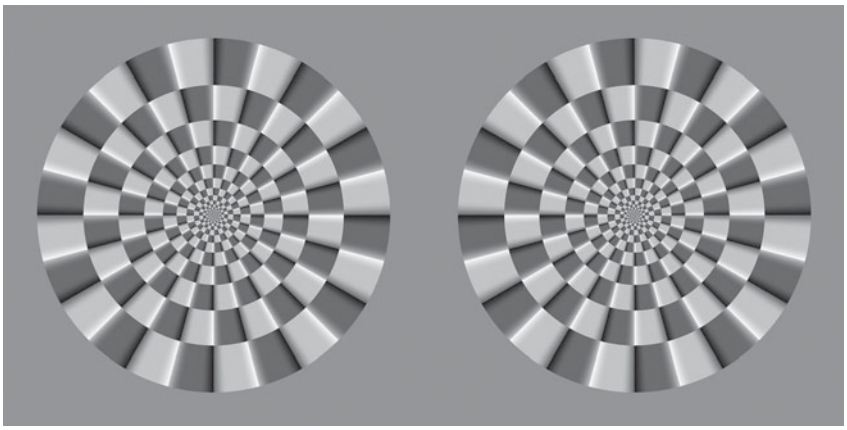


Figure IV.68-6. Continued

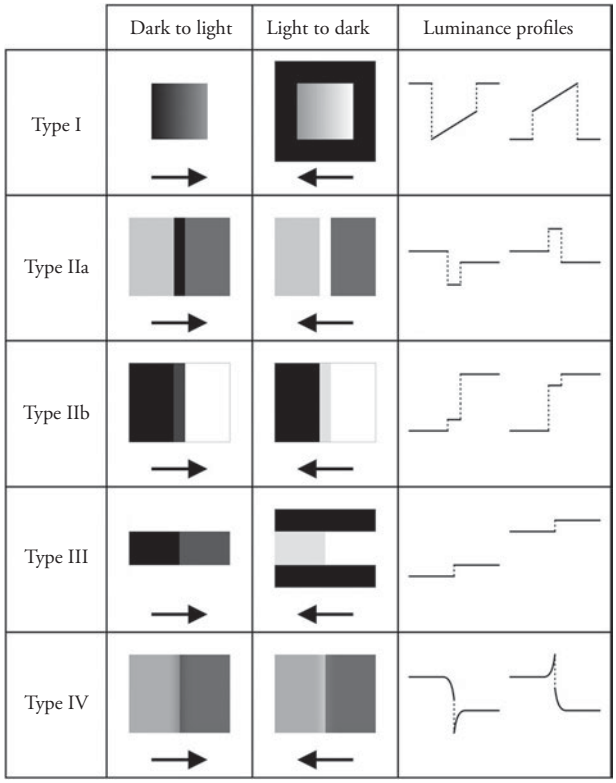


Figure IV.68-7. A classification of the "optimized Fraser-Wilcox illusion" family. Here are five types, each with two opposing illusions ("dark to light" and "light to dark"). The basic images of each type and their luminance profiles are shown. Arrows indicate the direction of illusory motion. (Figure by Akiyoshi Kitaoka.)

80s, saw the illusion. Moreover, about 1% of young participants did not see the illusion, while they did not report any accommodative problems or any other disorder in vision. In addition, it was suggested that infants ages 6 to 8 months see this illusion (Kanazawa, Kitaoka, & Yamaguchi, 2013).

The factor of race or culture might play some part in perception of the rotating snakes illusion. Billino et al. (2009; in Germany) examined 38 senior adults (65 to 82 years old), 23.7% of whom saw the illusion. On the other hand, my study (Japan) showed that 86.2% of 123 senior adults (65 to 82 years old) saw this illusion. The majority of the younger observers saw this illusion in both studies.

THE TIMING-DIFFERENCE MODEL AND THE EYE-MOVEMENT MODEL

The rotating snakes illusion is assumed to depend on timing differences between neural responses to different parts of the figure. Conway et al. (2005) explained this illusion in terms of contrast-dependent response timing differences and related it to four-stroke apparent motion (Anstis & Rogers, 1986; Mather & Murdoch 1999). Backus and Oruç (2005) also explained this illusion (referring to "repeated asymmetry patterns") in terms of response timing differences, while they also took into account the process of adaptation to luminance as well as contrast. According to these timing-difference models, illusory motion arises from peak shifts of perceived patterns.

On the other hand, Murakami et al. (2006) stressed the role of fixational eye movements in this illusion and a failure

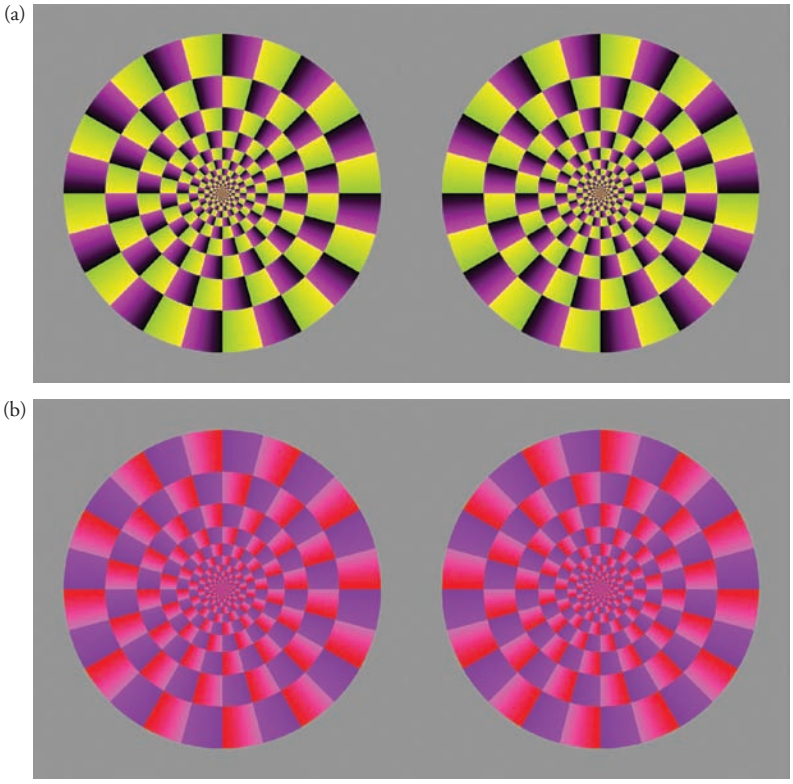


Figure IV.68-8. Color enhancement. In each panel, the left disk appears to rotate clockwise while the right one counterclockwise. (a) Red or blue (purple when mixed) enhance the illusion from dark to light shading while yellow or green increase the illusion from light to dark shading. (b) A particular combination of red and purple gives a strong illusion. (Figure by Akiyoshi Kitaoka.)

Table IV.68-1. The effect of age on the "Rotating snakes" illusion

Age	n	Score			
		0	1	2	3
0–9	169	0%	6%	17%	78%
10–19	710	1%	6%	19%	74%
20–29	219	1%	6%	27%	65%
30–39	219	1%	7%	21%	70%
40–49	221	2%	9%	29%	60%
50–59	138	8%	19%	30%	43%
60–69	98	5%	28%	38%	30%
70–79	61	15%	25%	33%	28%
80–92	28	21%	7%	39%	32%
Total	1863	3%	9%	24%	64%

Rating scores 0, 1, 2, 3 mean that observers reported no illusion, a slight illusion, an illusion, a strong illusion, respectively. Percentages indicate proportions in each age.

to compensate for the micromovements and explained it with the gradient model, which assumes a negatively biased temporal kernel. Relatedly, Fermüller et al. (2010) proposed the energy filter model assuming spatial symmetry and temporal asymmetry. According to these "eye movement" models, there is asymmetry in strength of motion signals depending on the direction of a moving pattern on the retina, and illusory motion arises from a difference of motion signals provided constantly by miniature eye movements. The involvement of fixational eye movement was supported by Kuriki et al. (2008) and Beer, Heckel, and Greenlee (2008).

Positive afterimages of the rotating snakes, given to dark-adapted eyes by a flash, did not generate the rotating snakes illusion (Murakami et al., 2006). This finding supports the eye-movement model. But the timing-difference model could also explain this finding by assuming that not a single flash but more frequent refreshments are necessary to maintain the illusion.

Using functional magnetic resonance imaging (fMRI) technology applied to the study of human brains, Kuriki et al. (2008) examined neural responses to the rotating snakes image (compared with responses to the control image) and revealed that neural activity of the motion-sensitive area hMT+ is correlated with the illusion. Moreover, this correlation was observed in two conditions that allowed eye movements but not in the condition that suppressed them. This finding supports the eye-movement model. Furthermore, using the fMRI adaptation technique, Ashida, Kuriki, Murakami, Hisakata, and Kitaoka (2012) revealed the activity of the cortical network for motion processing from V1 through hMT+.

RECENT STUDIES

Using an optimized Fraser-Wilcox illusion array, Beer et al. (2008) observed the interaction of arrays placed in the central and peripheral fields. They revealed that illusory motion was not perceived in the central field and occurred only in the peripheral field. The illusion was strongest when the array of the latter was opposed to that of the former (the "incongruent" condition). Moreover, given the central field, which has the array potentially giving the illusion, illusion in the "opposite" direction was induced to the peripheral field even if it did not have the array giving illusion. Furthermore, this induction was absent when the central field was physically moving.

In connection, Tomimatsu, Ito, Seno, and Sunaga (2010) showed that the rotating snakes illusion is weak when the image is physically moved and claimed that "the image should remain stationary (without being refreshed) for some time on the same retinal position (p. 721)." To support the timing-difference model in accordance with this observation, an additional assumption that "refreshment is suppressed during smooth pursuit" might be necessary. On the other hand, to support the eye-movement model in accordance with this observation, an additional assumption (e.g., "miniature eye movements are suppressed during smooth pursuit") would be necessary.

Hisakata and Murakami (2008) confirmed the anecdote that the rotating snakes illusion is better perceived in the visual periphery and is positively correlated with retinal luminance and suggested the involvement of a transient processing system by examining the human temporal impulse response. Cantor, Tahir, and Schor (2010) demonstrated that the rotating snakes illusion decreased when observers viewed the image through a pinhole or when they defocused it. They explained this finding by taking into account the modulation transfer function. Idesawa (2010) explained the optimized Fraser-Wilcox illusion Type IIa "by considering the isotropic spatial filtering process with time varying factor and visual reset (p. 557)." Tomimatsu, Ito, Sunaga, and Remijn (2011) showed evidence for the involvement of adaptation process both in the Fraser-Wilcox illusion and the optimized illusion.

In sum, many phenomena related to the Fraser-Wilcox illusion and its extension have been revealed, but their mechanisms are yet to be clearly understood.

ACKNOWLEDGMENTS

This study was supported by JSPS Grant-in-Aids for Scientific Research (A18203036 and A22243044) and a Ritsumeikan University R-GIRO program awarded to the author. Special gratitude is extended to Akemi Yamamoto, Nobuko Takamatsu, and Chiba Museum of Science and Industry for their help in collecting data for the study of age and illusion. I also thank Hiroshi Ashida and Arthur Shapiro for their comments.

REFERENCES

Anstis, S. M. (1970). Phi movement as a subtraction process. *Vision Research*, *10*(12), 1411–1430.

Anstis, S. (2000). Footsteps and inchworms: Illusions show that contrast affects apparent speed. *Perception*, *30*(7), 785–794.

Anstis, S. M., & Rogers, B. J. (1975). Illusory reversal of visual depth and movement during changes of contrast. *Vision Research*, *15*(8), 957–961.

Ashida, H., Kuriki, I., Murakami, I, Hisakata, R., & Kitaoka, A. (2012). Direction-specific fMRI adaptation reveals the visual cortical network underlying the "rotating snakes" illusion. *NeuroImage*, *61*(4), 1143–1152.

Backus, B. T., & Oruç, İ. (2005). Illusory motion from change over time in the response to contrast and luminance. *Journal of Vision*, *5*(11), 1055–1069.

Becker, C., & Mollon, J. D. (2007). Heritability of interindividual variation in illusory-motion perception: A study that exploits the Internet. *Perception*, *36*(ECVP Abstract Suppl.).

Beer, A. L., Heckel, A. H., & Greenlee, M. W. (2008). A motion illusion reveals mechanisms of perceptual stabilization. *PLoS ONE*, *3*(7), e2741.

Billino, J., Hamburger, K., & Gegenfurtner, K. R. (2009). Age effects on the perception of motion illusions. *Perception*, *38*(4), 508–521.

Cantor, C. R. L., Tahir, H. J., & Schor, C. M. (2010). Is the rotating snakes an *optical* illusion? *Journal of Vision*, *10*(7), 824a.

Cloete, S., Wallis, G., Walterso, L., & Broerseo, P. J. (2004). No evidence of perceptual dimorphism in the Fraser-Wilcox illusion. *Perception*, *33*(Suppl.), 153.

Conway, R. B., Kitaoka, A., Yazdanbakhsh, A., Pack, C. C., & Livingstone, M. S. (2005). Neural basis for a powerful static motion illusion. *Journal of Neuroscience*, *25*(23), 5651–5656.

Faubert, J., & Herbert, A. M. (1999). The peripheral drift illusion: A motion illusion in the visual periphery. *Perception*, *28*(5), 617–621.

Fermüller, C., Ji, H., & Kitaoka, A. (2010). Illusory motion due to causal time filtering. *Vision Research*, *50*(3), 315–329.

Fogel, D. B. (2002). In Memoriam: Alex S. Fraser (1923–2002). *IEEE Transactions on Evolutionary Computation*, *6*(5), 429–430.

Fraser, A. (2006). Alex Fraser, geneticist and painter. http://doctoralexfraser.blogspot.jp/

Fraser, A., & Wilcox, K. J. (1979). Perception of illusory movement. *Nature*, *281*(5732), 565–566.

Gregory, R. L., & Heard, P. F. (1983). Visual dissociations of movement, position, and stereo depth: Some phenomenal phenomena. *Quarterly Journal of Experimental Psychology*, *35A*, 217–237.

Helmholtz, H. (1962). *Treatise on physiological optics*. Vol. 2 (J. P. C. Southall, Trans.). New York: Dover. (Original work published 1867)

Hine, T. J., Cook, M., & Rogers, G. T. (1995) An illusion of relative motion dependent upon spatial frequency and orientation. *Vision Research*, *35*(22), 3093–3102.

Hisakata, R., & Murakami, I. (2008). The effects of eccentricity and retinal illuminance on the illusory motion seen in a stationary luminance gradient. *Vision Research*, *48*(19), 1940–1948.

Horn, B. (1975). Obtaining shape from shading information. In P. H. Winston (Ed.), *The psychology of computer vision* (pp. 115–155). New York: McGraw-Hill.

Idesawa, M. (2010). A model of illusory motion perception in still figures. *Optical Review*, *17*(6), 557–561.

Kanazawa, S., Kitaoka, A., & Yamaguchi, M. K. (2013). Infants see illusory motion in static figures. *Perception*, *42*(8), 828–834.

Kitaoka, A. (2006a). Configurational coincidence among six phenomena: A comment on van Lier and Csathó (2006). *Perception*, *35*(6), 799–806.

Kitaoka, A. (2006b). The effect of color on the optimized Fraser-Wilcox illusion. Paper presented for the 9th L'ORÉAL Art and Science of Color Prize, Tokyo. http://www.psy.ritsumei.ac.jp/~akitaoka/loreal2006paper.doc

Kitaoka, A. (2007, May). Phenomenal classification of the "optimized" Fraser-Wilcox illusion and the effect of color. Poster presented at DemoNight, VSS2007, GWiz, Sarasota, FL.

Kitaoka, A. (2008a, September). Optimized Fraser-Wilcox illusions: A pictorial classification by Akiyoshi Kitaoka. Paper presented at the 72nd Annual Convention of the Japanese Psychological Association. Hokkaido University, Sapporo, Japan. http://www.psy.ritsumei.ac.jp/~akitaoka/nisshin2008ws.html

Kitaoka, A. (2008b). *Trick eyes mechanism*. Tokyo: Kanzen. (in Japanese)

Kitaoka, A. (2010a). *Introduction to visual illusions*. Tokyo: Asakura-shoten. (in Japanese)

Kitaoka, A. (2010b). The Fraser illusion family and the corresponding motion illusions. *Perception*, *39*, (ECVP Abstract Suppl.), 178.

Kitaoka, A. (2012). The Fraser-Wilcox illusion and its extension. *Perception*, *41*(ECVP Abstract Suppl.), 91.

Kitaoka, A. (2015). Akiyoshi's illusion page. http://www.ritsumei.ac.jp/~akitaoka/index-e.html

Kitaoka, A., & Ashida, H. (2003). Phenomenal characteristics of the peripheral drift illusion. *VISION*, *15*(4), 261–262.

Kitaoka, A, & Ashida, H. (2004, January). A new anomalous motion illusion: The "central drift illusion." Paper presented at the winter meeting of the Vision Society of Japan. Kogakuin University, Tokyo. http://www.psy.ritsumei.ac.jp/~akitaoka/VSJ04w.html

Kitaoka, A., & Ashida, H. (2007). A variant of the anomalous motion illusion based upon contrast and visual latency. *Perception*, *36*(7), 1019–1035.

Kuriki, I., Ashida, H., Murakami, I., & Kitaoka, A. (2008). Functional brain imaging of the rotating snakes illusion by fMRI. *Journal of Vision*, *8*(10):16, 1–10.

Leviant, I. (1996). Does "brain-power" make Enigma spin? *Proceedings of Royal Society of London, B*, *263*(1373), 997–1001.

MacKay, D. M. (1957). Moving visual images produced by regular stationary patterns. *Nature*, *180*(4591), 849–850.

Mather, G., & Murdoch, L. (1999). Second-order processing of four-stroke apparent motion. *Vision Research*, *39*(10), 1795–1802.

Mather, G., Verstraten, F., & Anstis, S. (1998). *The motion aftereffect: A modern perspective*. Cambridge, MA: MIT Press.

Murakami, I., Kitaoka, A., & Ashida, H. (2006). A positive correlation between fixation instability and the strength of illusory motion in a static display. *Vision Research*, *46*(15), 2421–2431.

Naor-Raz, G., & Sekuler, R. (2000). Perceptual dimorphism in visual motion from stationary patterns. *Perception*, *29*(3), 325–335.

Pinna, B., & Brelstaff, G. J. (2000). A new visual illusion of relative motion. *Vision Research*, *40*(16), 2091–2096.

Pinna, B., & Spillmann, L. (2002). A new illusion of floating motion in depth. *Perception*, *31*(12), 1501–1502.

Ramachandran, V. S. (1988). Perception of shape from shading. *Nature*, *331*(6152), 163–166.

Shapiro, A. G., Knight, E. J., Palmor, Y., & Mancusi, D. (2006). Swimmers, eels and other gradient-gradient illusions. Exhibits for Vision Sciences Society Demonstration Night, Sarasota, FL. http://www.shapirolab.net/

Spillmann, L. (2013). The Ōuchi–Spillmann illusion revisited. *Perception*, *42*(4), 413–429.

Spillmann, L., Heitger, F., & Schuller, S. (1986). Apparent displacement and phase unlocking in checkerboard patterns. Paper presented at the 9th European Conference on Visual Perception, Bad Nauheim, Germany.

Tomimatsu, E., Ito, H., Seno, T., & Sunaga, S. (2010). The "rotating snakes" in smooth motion do not appear to rotate. *Perception*, *39*(5), 721–724.

Tomimatsu E., Ito, H., Sunaga, S., & Remijn, G. B. (2011). Halt and recovery of illusory motion perception from peripherally viewed static images. *Attention, Perception, & Psychophysics*, *73*(6), 1823–1832.

Zanker, J. M., Doyle, M., & Walker, R. (2003). Gaze stability of observers watching op art pictures. *Perception*, *32*(9), 1037–1049.

Chapter 69
Induced Motion

Jasmin Léveillé and Arash Yazdanbakhsh

HISTORY AND PHENOMENAL DESCRIPTION

The motion of one object can sometimes make another, usually nearby object appear to move in the opposite direction. This phenomenon, termed induced motion, is perhaps one of the most studied demonstrations of the notion of object-centric perception, according to which the visual system is geared to perceive the relative configuration of objects in a scene rather than their configuration relative to the observer.

That the motion of one object can impart opposite-directed motion to another in a natural setting has been noted several times throughout history. In his book *Optics*, Ptolemy (c. 160–170 ad) described how a boat that stands still in the flow of a river appears to be moving upriver whereas the water appears to be standing still (Smith, 1996). A more mathematical description of induced motion had already been provided by Euclid (c. 300 bc). Perhaps the easiest way to demonstrate induced motion in a natural setting is to consider what happens when looking at the moon on a cloudy and windy night. Here the drifting motion of the clouds in one direction can clearly make the moon appear to be moving in the other direction. The polymath Ibn al-Haytham ("Alhazen") wrote on this phenomenon in the first half of the 11th century (Howard, 1996; Sabra, 1989). Similar phenomena were reported several times in the 18th century as well (Wade, 2000; Wade & Swanston, 1987).

More recently, a wide variety of controlled displays have been designed to study induced motion in a laboratory setting. The most well-known version was introduced in a demonstration due to Duncker (1938). In Duncker's display, a static dot (the target) is surrounded by a frame (the inducer) that moves on the frontoparallel plane (Fig. IV.69-1a). When an observer sits in front of such a display in a dimly lit environment, the dot is seen to be moving in a direction opposite to the real motion of the frame. Target motion can be perceived irrespective of whether the frame's speed is below or above the detection threshold (Wallach et al., 1978). The magnitude of induced motion perceived can easily be measured by such procedures as having observers match the length of the perceived displacement of the dot (e.g., Brooks & Sherrick, 1994), by finger-pointing to the perceived location reached by the target after a stimulus trial (e.g., Post et al., 2008), or by verbally reporting on the presence, absence, or direction of motion (Di Vita & Rock, 1997; Murakami, 1999).

Wallach et al. (1978) proposed a variant of Duncker's induced-motion stimulus in which the target dot is also endowed with a real motion component along the vertical axis (Fig. IV.69-1b). In such a display, the target dot appears to be moving diagonally, as if its motion combined two components: the actual, vertical component of the dot and the induced, horizontal component. The "amount" of induced motion perceived is then directly related to the tilt of the target dot's trajectory: the more horizontal the trajectory, the stronger the induced motion. A more accurate way of describing induced motion—suggested by Wallach et al.'s display—is thus to say that the reference frame induces a horizontal "component" of motion opposite to its own motion. This illusory motion component can in turn be combined with others (e.g., the real vertical component of the dot) to yield the final percept of the dot moving along a given trajectory. A commonly used method to measure the amount of induced motion in displays inspired from Wallach et al.'s study is to use a rotatable bar as a probe. Participants adjust the tilt of this probe while free-viewing the stimulus until it matches the tilted trajectory of the target dot (e.g., Gogel & MacCracken, 1979). A straightforward trigonometric transformation is then optionally applied to recover the horizontal component (the component that is induced) from the tilt angle.

A considerable amount of research has focused on identifying which conditions promote induced motion and which do not. For example, a reasonable hypothesis might be that the target element should "belong" to the inducer. In the stimuli of Figure IV.69-1, for example, the dot belongs to the frame by virtue of being enclosed within the rectangular frame. In such an arrangement, the visual system would "deduce" that the motion of the dot should be affected by that of the frame—more than, say, if the dot were not surrounded by the frame. Note that, in Duncker's stimulus, for the dot to appear to be static on the monitor it must be truly moving in a direction opposite to that of the frame, hence the visual system infers an induced component (Di Vita & Rock, 1997). Unfortunately, belongingness—at least when based on enclosure of the dot by the frame—is not necessary: it is easy to obtain induced motion even in displays where the target is not enclosed in the frame but is adjacent to it, in which case it is harder to say that the dot

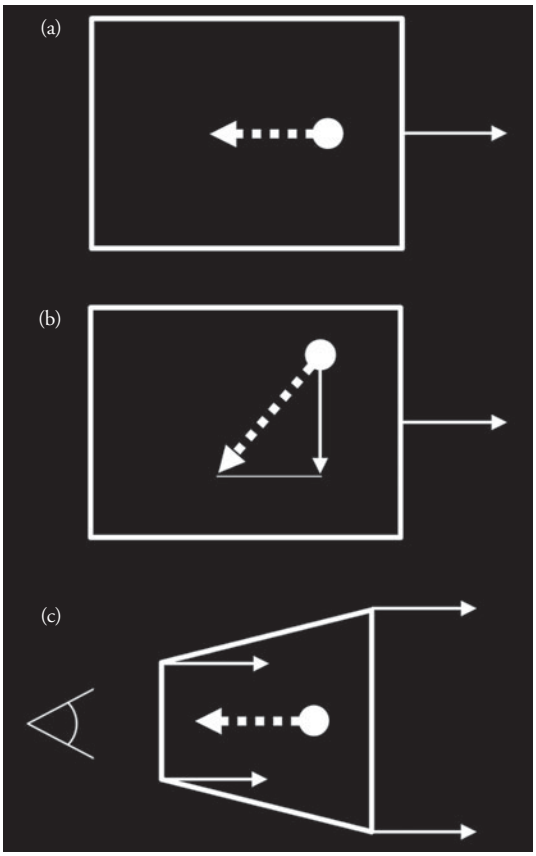


Figure IV.69-1. Induced motion stimulus displays. Thin and thick, dashed arrows stand for real and induced motion components, respectively. (a) Duncker's induced motion display consists of a static dot surrounded by a moving frame. The induced motion percept is that of a dot moving in the opposite direction. (b) By comparison, the target dot in the display of Wallach et al. (1978) has a real vertical motion component that, when combined with the induced component, gives rise to a diagonal motion percept. In the actual stimulus display of Wallach et al., vertical bars were also on the surface of the inducer frame. (c) Sample display for induced motion in depth. The eye of the observer is denoted here by the arc on the left. A frame moving away from the observer makes the target dot appear to move toward the observer.

belongs to the frame (Gogel & Koslow, 1972). Yet it is also clear that induced motion works only up to a certain amount of separation between frame and target, suggesting the involvement of a form of spatially local computation. The nature of the inducer frame also affects the induced-motion component. For example, frames composed of a number of distinct elements—rather than a single rectangle—appear to exert a greater influence on the target (Day et al., 1979). Using transparent motion inducers (e.g., a cloud of coherently moving dots), the luminance contrast of the inducer can determine whether induced motion or motion assimilation occurs (Murakami, 1999).

By far most experiments to date have focused on induced motion with stimulus elements whose motions are restricted to the frontal plane. Given that it works on that plane, it is natural to wonder whether an analogous phenomenon emerges when stimulus elements move in depth. A few experiments show that, when using an inducer that oscillates in depth, induced motion occurs in that dimension: for example, if a frame is seen as moving toward the observer, the target dot appears to move away and vice versa (Fig. IV.69-1c). There is debate as to whether induced motion in depth is as compelling as when it is along the frontal plane and how conducive different depth cues are for the emergence of an induced component (Nefs & Harris, 2008). Nevertheless, the mere possibility of induced motion in depth is enough to suggest that encoding relative motion is a general strategy employed by the visual system rather than only a byproduct of, say, opponency between motion-tuned neurons. The link between the three-dimensional (3D) structure of visual space and induced motion is further characterized by the observation that separating, in different depth planes, a target from an inducer moving along the frontal plane does not affect frontal induced motion (Di Vita & Rock, 1997; Léveillé & Yazdanbakhsh, 2010).

CURRENT HYPOTHESES

Relatively little conceptual progress has been made since Duncker's (1938) original induced-motion experiment. Most attempts at explaining induced motion refer to object-centric (or object-centered, or pattern-centric) reference frames.

To understand this idea requires delving a little deeper into the notion of coordinate frames. Different coordinate systems are generally thought to be involved in visual perception (Wade & Swanston, 1996). At one extreme, in a retinocentric system, position is encoded relative to the location where the light reflected by an object impinges on the retina. Assuming a stable fixation point and a well-defined origin on the retina, the trajectory of a moving object would thus be uniquely encoded as a two-dimensional vector. Although this scheme is perhaps the simplest way to encode position, such a system would fail to exploit much of the higher-order information present in natural scenes. At the other extreme, in an allocentric reference frame, object coordinates are encoded relative to some fixed reference in the 3D world. According to this scheme, the motion of an object would be uniquely represented as a 3D vector with the origin somewhere in the environment and would be invariant to various changes in viewpoint. Although this scheme may naturally provide a simple means to process interacting objects located at different locations, it is arguably the most difficult to establish as it requires multiple successive transformations (Wade & Swanston, 1996). Induced motion may be framed as the manifestation of yet another type of coordinate frame, one in which the motion of an object is encoded relative to the motion (or position) of another. Although consistent with available evidence, this hypothesis leaves unanswered a number of questions (see Brosgole, 1968, for an alternative viewpoint). For example, what are the benefits of an object-centric encoding? On what basis is one stimulus element assigned to be the reference frame for another? With respect to this last question, larger stimulus elements are usually the ones to induce motion in smaller ones (Rock, 1975).

In terms of the actual computations that may yield induced motion, one notion that may be of interest is that of vector subtraction. According to vector subtraction, the motion vector corresponding to the reference frame is

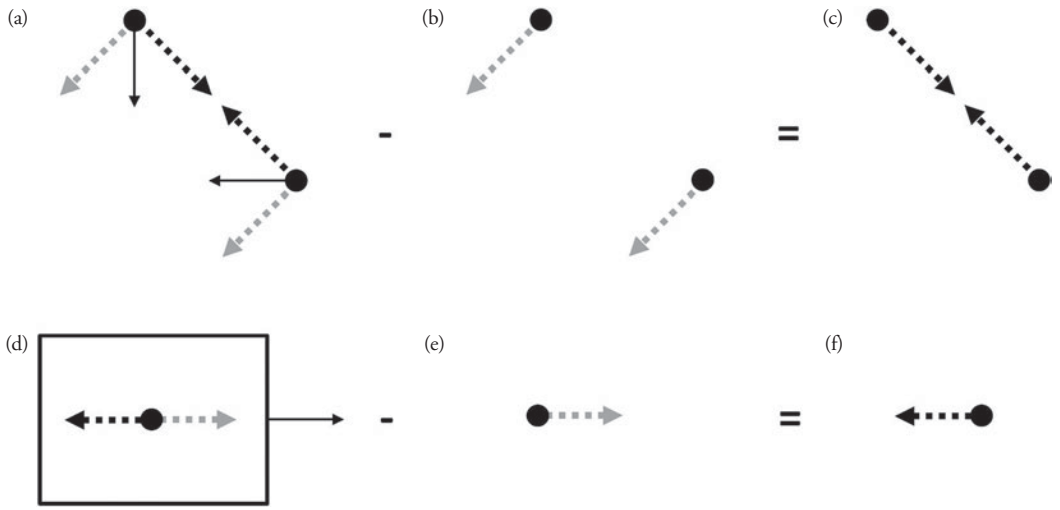


Figure IV.69-2. Induced motion, vector decomposition, and vector subtraction. Real motions are indicated by thin arrows, the motion of the reference frame by dashed gray arrows, and the induced component by dashed black arrows. (a) through (c) shows one of the stimuli studied by Johansson (1950) in his investigations of vector decomposition. (a) The display consists of two dots oscillating toward a common point at the lower left. The motion component that is common to both dots in the pair (dashed gray arrows) defines the reference frame's motion. When subtracting this component (b), the remaining component is that of the dots moving toward each other (c), as reported by observers. It was also demonstrated that the component motion perception can be altered during smooth pursuit eye movements (Festinger et al., 1976; Johansson, 1950, p. 89). (d) through (f) shows an application of vector subtraction to induced motion. (d) Here the reference frame's motion component coincides with the real, rightward motion of the rectangle. When subtracted (e), the remaining component (f) corresponds to induced motion.

somehow subtracted from that of the target—alternatively, the negative of the frame's vector is added. Straightforward geometrical considerations show that the perceived motion resulting from such subtraction corresponds indeed to the remaining motion vector. From this point of view, induced motion may well be a close analog of the phenomenon of vector decomposition (Johansson, 1950). In a vector decomposition stimulus, the motion component that is common to all elements of a stimulus (e.g., a pair of moving dots) is subtracted, yielding relative motion components and common motion components as distinctly perceived motions (Fig. IV.69-2). The same notion of vector subtraction therefore applies here but using common motion instead of reference frame motion. The close link with Duncker's induced motion was in fact already acknowledged in early work on vector decomposition (Johansson, 1994).

It should be noted that the fact that the induced component and the actual motion of the frame can be simultaneously perceived is puzzling from the viewpoint of a strict vector analysis. If the motion of the dot is perceived relative to the frame, and the motion of the frame is perceived relative to the background (or to any other object that may act as its reference frame), why is the visual system unable to integrate the two motions to correctly encode the dot's absolute motion? To resolve this conundrum, Duncker (1938) introduced the notion of *separation of systems*. Accordingly, the motion of the dot relative to the frame is unaffected by the relationship of the motion of the frame to the rest of the visual scene (it is not a transitive operation). Whether such separation of systems holds in a more general setting, and how such separation may be implemented by neural circuits, remains unclear.

Center-surround antagonism of motion detectors is typically invoked as the potential neural mechanism of induced motion (Murakami, 1999). Figure IV.69-3 shows an example of a particular center-surround configuration that could

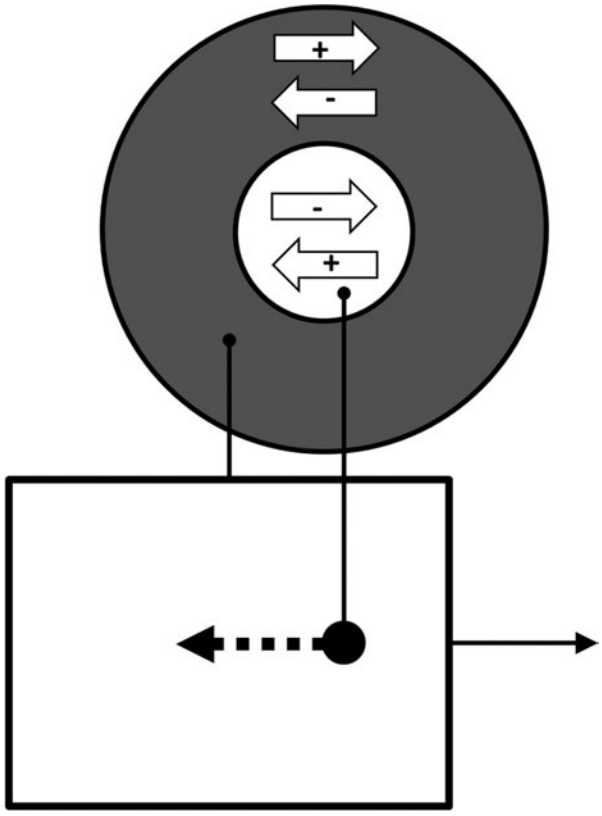


Figure IV.69-3. One example of motion opponency in induced motion. The center-surround diagram at the top represents a double-opponent, motion-receptive field for a unit tuned to leftward motion in its center. A neuron with this kind of receptive field is excited either by rightward motion in its surround or leftward motion in its center and is inhibited by either leftward motion in its surround of rightward motion in its center. Such a unit located in the vicinity of the target dot would be strongly activated by the surround, rightward motion of the frame and hence would trigger a percept of leftward motion.

be expected to implement an operation akin to vector subtraction. Although this mechanism would seem to account for the emergence of motion (in static target displays) or for the change in direction (in moving target dot displays), it is not clear how it would handle the influence of the speed of elements. This last concern stems from the considerations that many motion-detector cells are tuned to speed in addition to direction, while vector subtraction provides a good description of data on induced motion at a wide range of inducer and target speeds (Léveillé & Yazdanbakhsh, 2011).

Virtually no work has been done addressing how neural circuits implement induced motion. Given that the phenomenon involves both sophisticated aspects of form and motion, it is unlikely that center-surround competition among motion detectors will be sufficient to explain induced motion.

CONCLUSIONS

Induced motion exemplifies the notion of object-centric perception. Relative to a moving frame, the motion of an object truly contains a component not present in its real motion. From the viewpoint of relative motion then, there is no illusion in the phenomenon of induced motion. The illusion instead stems from the fact that, while the visual systems seems to process the induced motion stimulus in an object-centric way, the induced component is mistakenly perceived as if it were really in an egocentric reference frame.

REFERENCES

Brooks, J. N., & Sherrick, M. F. (1994). Induced motion and the visual vertical: Effects of frame size. *Perceptual and Motor Skills*, *79*, 1442–1450.

Brosgole, L. (1968). An analysis of induced motion. *Acta Psychologica*, *28*, 1–44.

Day, R. H., Miller, J., & Dickinson, R. G. (1979). Induced movement as nonveridical resolution of displacement ambiguity: Effects of enclosure and number of field elements. *Perception & Psychophysics*, *25*, 23–28.

Di Vita, J. C., & Rock, I. (1997). A belongingness principle of motion perception. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 1343–1352.

Duncker, K. (1938). Uber induzierte Bewegung (Ein Beitrag zur Theorie optisch wahrgenommener Bewegung) [Concerning induced movement]. In W.D. Ellis (Ed. And Trans.), *Source book of Gestalt psychology* (pp.161–172). London: Routledge Kegan Paul. (Reprinted from Psychologische Forschung, 1929, 12, 180–259.

Festinger, L., Sedgwick, H. A. & Holtzmann, J. D. (1976). Visual perception during smooth pursuit eye movements. *Vision Research*, *16*, 1377–1386.

Gogel, W. C., & Koslow, M. (1972). The adjacency principle and induced movement. *Perception & Psychophysics*, *11*, 309–314.

Gogel, W. C., & MacCracken, P. J. (1979). Depth adjacency and induced motion. *Perceptual and Motor Skills*, *48*, 343–350.

Howard, I. P. (1996). Alhazen's neglected discoveries of visual phenomena. *Perception*, *25*, 1202–1217.

Johansson, G. (1950). *Configurations in motion perception*. Uppsala, Sweden: Almqvist & Wiksell.

Johansson, G. (1994). A practical theorist: An interview with William Epstein, edited by Sheena Rogers. In G. Jansson, S. S. Bergstrom, & W. Epstein (Eds.), *Perceiving events and objects* (pp. 3–28). Mahwah, NJ: Lawrence Erlbaum.

Léveillé, J., & Yazdanbakhsh, A. (2010). Speed, more than depth, determines the strength of induced motion. *Journal of Vision*, *10*, 1–9.

Léveillé, J., & Yazdanbakhsh (2011). Vector subtraction in velocity space. [Abstract] 15th International Conference on Cognitive and Neural Systems, Boston, MA.

Murakami, I. (1999). Motion-transparent inducers have different effects on induced motion and motion capture. *Vision Research*, *39*, 1671–1681.

Nefs, H. T., & Harris, J. M. (2008). Induced motion in depth and the effects of vergence eye movements. *Journal of Vision*, *8*, 1–8.

Post, R. B., Welch, R. B., & Whitney, D. (2008). Egocentric and allocentric localization during induced motion. *Experimental Brain Research*, *191*, 495–504.

Rock, I. (1975). *An introduction to perception*. New York: Macmillan.

Sabra, A. I. (1989). *The optics of Ibn Al-Haytham. Books I–III: On direct vision*. London: Warburg Institute.

Smith, M. A. (1996). *Ptolemy's theory of visual perception: An English translation of the Optics with introduction and commentary*. Transactions of the American Philosophical Society, New Series, 86(2). Philadelphia: American Philosophical Society.

Wade, N. J. (2000). *A natural history of vision*. Cambridge, MA: MIT Press.

Wade, N. J., & Swanston, M. T. (1987). The representation of non-uniform motion: Induced movement. *Perception*, *16*, 555–571.

Wade, N. J., & Swanston, M. T. (1996). A general model for the perception of space and motion. *Perception*, *25*, 187–194.

Wallach, H., Bacon, J., & Schulman, P. (1978). Adaptation in motion perception: Alteration of induced motion. *Perception & Psychophysics*, *24*, 509–514.

Chapter 70
The Freezing Rotation Illusion

Erika N. Lorincz and Max R. Dürsteler

DESCRIPTION OF THE FREEZING ROTATION ILLUSION

When an object such as a picture of an airplane (Fig. IV.70-1a) or a textured disk ("center," Fig. IV.70-1b) is placed over a patterned background ("surround"), singular phenomena can be perceived depending on the relative rotations of the center and of the surround. Both center and surround rotate around the same axis perpendicular to the display. An unidirectional constant velocity rotation is applied to the center, while the surround rotates sinusoidally back and forth.

The freezing rotation (FRI) refers to the illusion where the center is perceived as slowing down or even coming to a standstill. This illusion occurs when the center and the surround rotate in the same direction and the center's angular velocity is much lower than the surround's maximal velocity (Fig. IV.70-1b, left). By contrast, the opposite effect arises (i.e., the center is perceived as rotating faster than its physical angular velocity) when the center and surround rotate in opposite directions (Fig. IV.70-1b, right; see interactive figure IV.70-1 for both illusions). These main effects are accompanied by other phenomena depending on the relative rotational profiles and velocities of the center and the surround, which we attempt to explain here.

SIGNIFICANCE OF THE FREEZING ROTATION ILLUSION

Motion detection by the visual system serves both our stabilization in a possibly shaking world and the recognition of a moving object. However, a dual-purpose visual system is unlikely to exist without incurring some compromises that give rise to artifacts (Anstis & Reinhardt-Rutland, 1976; Ito, 2012). The FRI is such an example, where our visual system makes some inferences about the perceived rotation of an enclosed object from the rotations of the surround.

The stimulus configuration in the original FRI with a circular center rotating on a larger oscillating surround (Dürsteler, 2005, 2008) is very close to the stimulus configuration used by Duncker (1929) in his classical experiments regarding induced rotational motion. He described that a stationary central disc mounted in front of a larger rotating

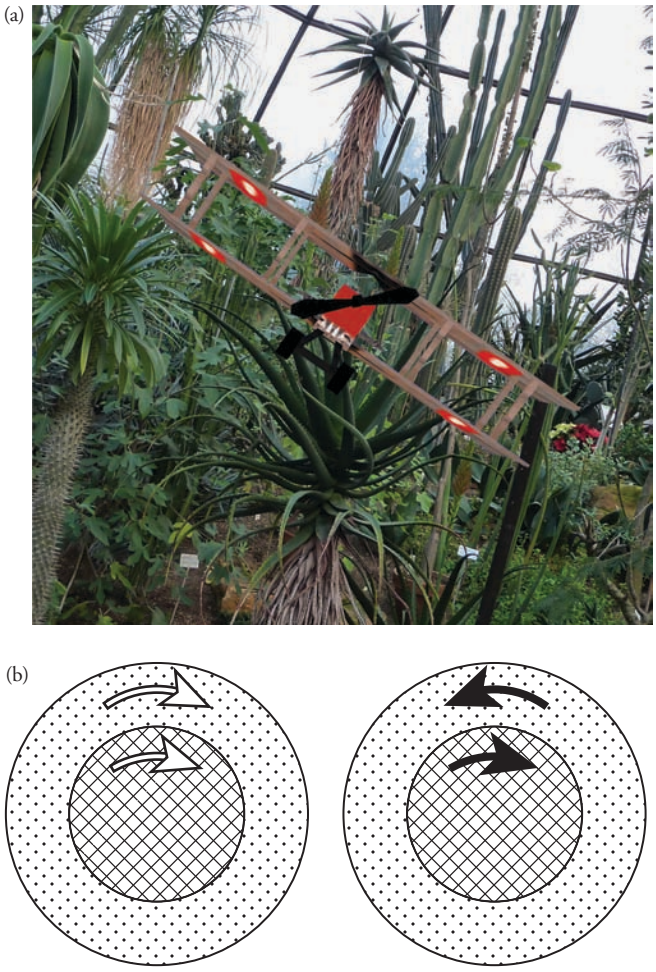


Figure IV.70-1. Representation of the stimulus composition for the freezing rotation illusion. (a) Screenshot of the accompanying application. (b) The arrows represent the rotation directions of the center and the surround. When the center and the surround turn in the same direction (left panel, light gray arrows), the center appears to slow down. When the center and the surround rotate in opposite directions (right panel, black arrows), the center appears to speed up. (Figure by Max R. Dürsteler.)

textured disc (see Table IV.70-1) appeared to rotate in the opposite direction to that of the disk behind. He called this phenomenon induced motion. Similarly to the FRI, the perceived counterrotation of the inner disk could be canceled by rotating the central disc by about 5.6° in the same direction as the surrounding disk. Using similar stimuli and procedures, Day (1981) confirmed the occurrence of induced motion and reported that its occurrence was neither affected by the angular velocity of the surround (counter to Duncker's description) nor by a gray gap between the center and the surround. By contrast, the perceived velocity of the center was enhanced with increasing spatial frequencies of the stimulus. In addition, rotational induced motion depends on the configuration of the display, with the peripheral location of the inducer (the rotating annulus) surrounding the induced static element, an important factor for eliciting motion of the central stimulus (Day, 1981; Reinhardt-Rutland, 1981). However, the area per se of the enclosing inducer can be reduced to a narrow annulus. Surprisingly adaptation to induced motion, as real movements, can elicit negative-motion aftereffects of the central static stimulus (i.e., an illusionary motion generated after disappearance of the stimulation, which goes in the opposite direction to that of the initially induced motion; Reinhardt-Rutland, 1981).

At first glance, the FRI might resemble induced motion. Dürsteler (2007, 2008) proposed that the illusion could be provoked by the induced motion resulting from the relative motion of the center and the surround. To check this, he tested whether a fixed constant velocity applied to the surround could influence the perceived velocity of a constant velocity rotating center (experiment a) and whether the reversed was also true (b). After a short presentation of the test stimulus (500 ms), participants were asked to judge whether either the center (a) or the surround (b) was perceived as rotating faster or slower than a probe stimulus. These experiments showed that only the surround's rotation influenced the perceived center's rotation (and not the opposite). But, counter to expectations, these results demonstrated an overestimation of the perceived angular velocity of the center for opposite rotation directions but not the FRI for center and surround corotations.

Wertheim and Paffen (2009) proposed that if induced motion were responsible for the effect, its magnitude should vary according to the velocities' difference between the center and the surround and so should be the strength of the illusion. Consequently they investigated the effect of center-surround relative motion on the occurrence of the FRI using different surround velocities in addition to varying the center velocity. During corotations, they observed the freezing effect only for very low centers' velocities but reported a fixed underestimation for higher ones, regardless of the amount of relative motion between center and surround. They concluded that induced motion could not possibly account for the illusion and that configural factors associated with the proper rotation direction and velocity conditions might cause the illusion to occur.

Induced motion is a common phenomenon that can also be observed during linear displacement and motion in depth (see, for review, Reinhardt-Rutland, 1988). However, its occurrence and magnitude seem to depend on complex interactions between parameters, among which luminance contrast, stimulus size, configuration, and appearance of the display (e.g., spatial frequency, shape, etc.) seem to play crucial roles (Murakami & Shimojo, 1996). Steady, unidirectional inducing stimulation produces less induced movement at higher speeds than oscillation (Becklen & Wallach, 1985; Wallach & Becklen, 1983). Those factors affecting induced motion could explain the discrepancy in the effects observed by Wertheim and Paffen (2009) and by Dürsteler (2007, 2008) and participate in eliciting the FRI.

Indeed, mechanisms to detect differential motion could be common to several types of subjective movements, such as induced motion, motion aftereffects, and motion capture (Murakami & Shimojo, 1996). Motion capture normally occurs when coherent motion of low-frequency elements (e.g., grating) influences the perceived position or displacement of a high-frequency stimulus (e.g., randomdots) to be in the same direction as the low-frequency elements (Ramachandran & Cavanagh, 1987). Murakami and Shimojo (1993, 1996) showed that motion capture can change to induced motion depending on the interaction between the eccentricity, the size, and the luminance contrast of the central area. The FRI could occur following a similar process.

Finally, similarly to induced motion, the question arises about the reference frame in which the surround and the center velocities are derived to provoke the perceived alteration of motion observed in the FRI (i.e., space centered, object centered, or texture centered; Reinhardt-Rutland, 1988). Counter to translational displacement, the "object" composed of the center and the background is stationary relative to space; only its texture is animated. However, the current illusion affects global motion, because the center is seen as moving relative to the surround but frozen relative to the external space (Wertheim & Paffen, 2009). Local displacements between center and surround that remain visible could have contributed to build a coherent motion percept (Lee & Lu, 2010).

In the following we present the FRI as well as its variants and try to provide explanations to these puzzling phenomena. The FRI is more than a trivial illusion, and trying to understand such phenomena might help unravel how the human visual system works.

VARIANTS AND RELEVANT PARAMETERS OF THE FREEZING ROTATION ILLUSION

The FRI as well as the related illusions we present here occur for particular conditions of relative rotation directions and velocities of the center and the surround. They are best elicited in central vision, although slightly different but interesting phenomena also occur with more peripheral vision. All can be elicited using the accompanying interactive figure IV.70-1, where one can choose different items for the center and the surround and change different stimulation parameters. The

Table IV.70-1. Suggested settings for the FRI and its variants. v = continuous rotation velocity; v_{max} = maximum speed; f = frequency; C = center; S = surround. The suggested surround's viewing angle is 15°.

	$Type_C$	v_C °/sec	v_{maxC} °/sec	f_C Hz	$Size_C$ %	$Opacity_C$ %	$Type_S$	v_S °/sec	v_{maxS} °/sec	f_S Hz	$Size_S$ %	$Opacity_S$ %
a	Airplane	12	0	0	60	100	Greenhouse	0	43.7	0.23	100	100
b	Airplane	12	0	0	60	100	Greenhouse	0	43.7	0.23	100	0
c	Plants	12	0	0	60	100	Greenhouse	0	43.7	0.23	70	100
d	Airplane	12	0	0	70	100	Greenhouse	0	43.7	0.23	23	100
e	Plants	12	0	0	70	50	Greenhouse	0	43.7	0.23	100	100
f	Airplane	>12~32	0	0	70	100	Greenhouse	0	43.7	0.23	100	100
g	Airplane	>32~47	0	0	70	100	Greenhouse	0	43.7	0.23	100	100
h	Airplane	>47	0	0	70	100	Greenhouse	0	43.7	0.23	100	100
i	Ran. dots	47	0	0	70	100	Ran. dots	0	43.7	0.23	100	100
j_1	Airplane	12	0	0	70	100	Greenhouse	–27.8	0	0	100	100
j_2	Airplane	12	0	0	70	100	Greenhouse	27.8	0	0	100	100
k	Plants	0	43.7	0.23	70	100	Greenhouse	12	0	0	100	100
l	Airplane	0	0	0	70	100	Greenhouse	12	0	0	100	100

stimulation parameters needed for variants (a) to (k) are listed in Table IV.70-1.

(a) The recommended parameters allow seeing the original FRI (see previous section for the complete description).
(b) Setting the opacity of the surround to zero allows one to check that the center is actually rotating in a smooth and continuous fashion.
(c) The FRI is still robust as long as even a small ribbon of the surround's perimeter is visible.
(d) It is not necessary that the entire central region be confined within the surround to elicit the FRI. Here the airplane is still perceived as intermittently slowing down, despite the fact that most of its wings and undercarriage overlaps the surround. If we further decrease the size of surround, the illusion disappears.
(e) It is not necessary that the center be spatially separated from its surround. Even if we see the "surround" oscillating underneath the rotating center, the intermittent freezing is still obvious.
(f) The FRI is progressively replaced by a perceived slowdown of the center, when the airplane's velocity is increased from 12 to about 32°/s.
(g) If we raise the center's velocity from 32°/s to about 47°/s, the center is perceived as "sticking" to it; that is, center and surround are perceived as rotating at the same velocity when they rotate in the same direction. The same percept as for the lower center's velocities is experienced for counterrotations.
(h) If the velocity of the central texture is even higher (>47°/s), the center is perceived again as rotating more slowly than its physical velocity during corotations and at the same velocity or faster during counterrotations.
(i) During corotation, we can make the apparent border between center and surround disappear using the same texture. Reliable border recognition requires that the ratio of angular velocities between textures remain above or below a certain threshold.
(j) Unidirectional rotation of the surround at constant velocity can also elicit most of these effects, that is, either an increase or a decrease of the perceived velocity of the center depending on whether it counter- (j_1) or corotates (j_2) with the surround.
(k) Applying a sinusoidal motion to the center (which increases the illusion when applied to the surround) and a continuous rotation to the surround does not modify the perceived velocity of the surround. In other words, while the surround's velocity influences the perceived motion of the center, the motion of the center never influences the perceived velocity of the surround.

The FRI and its variants demonstrate a profound influence of the surround rotation on the perceived angular velocity of the center. This idea is reinforced by the fact that when both center and surround rotate at constant velocities, only the perceived velocity of the center is altered.

DISCUSSION

During corotations of the center and the surround, the FRI varies as the central velocity is increased, from a total freeze to a slowdown of the center. By contrast, for counterrotations identical phenomena are observed regardless of the center's velocity (i.e., either no change or a perceived increase of the center's velocity).

Critical Parameters for the Freezing Rotation Illusion

When using the airplane and the greenhouse, the motionless percept of the rotating center is best observed for the center's velocities that equal about half the *average* velocity of the surround. If we consider that

(1) $v_{Avg\ Surround} \cong k \times v_{Center}$ where $v_{Avg\ Surround}$ is the surround's average angular velocity; v_{Center} the center's constant angular velocity; k the constant relative to the optimal surround-center ratio (here 2)

(2) $v_{Max\ Surround} \cong v_{Avg\ Surround} \times \frac{\pi}{2}$ where $v_{Avg\ Surround}$ is the surround's maximal angular velocity

(3) $\omega_f = \frac{v_{Max\ Surround}}{A}$ where ω_f is the surround's rotational speed, and A its amplitude

(4) $f_{Surround} = \frac{\omega_f}{2 \times \pi} = \frac{v_{Max\ Surround}}{2 \times \pi \times A}$ where f_{Surround} is the surround angular frequency. The optimal surround's maximal angular velocity to obtain the freezing effect could be adjusted by a parameter b that compensates for some bias (e.g., induced motion):

(5) $v_{Max\ Surround} = \frac{k \times \pi \times v_{Center}}{2} + b$ (°/s).

(6) Consequently $f_{Surround} = \frac{k \times \pi \times v_{Center} + 2 \times b}{2 \times \pi \times A}$ (Hz)

According to our prediction for the current display, and for simplification purposes, we set A = 0.53 rad ≈ 30° and b = 6°/s as constants (Table IV.70-2). This formula seems to work at least until a center velocity of 70°/s for sinusoidal stimulations (Table IV.70-2a).

This observation is to be taken with caution because controlled measurements were not made; the current freezing effect seems stronger, more consistent, and to cover a wider range of velocities than what was previously reported (Dürsteler, 2008; Wertheim & Paffen, 2009). How can these discrepancies be explained?

Wertheim and Paffen (2009) found the freezing effect to occur only for very low center velocities, while at higher ones it was underestimated only by a fixed amount (i.e., about 6.5°/s) regardless of the relative velocities between surround and center. Rather than their difference, which is usually used for translational motion (Wertheim & Reymond, 2007), the ratio k between surround and center velocities may be a better predictor for the occurrence of a motionless percept. For Wertheim and Paffen's data we estimated this ratio to be about 5.27 (instead of 2 for the airplane and the greenhouse; see Table IV.70-2b). For constant velocity rotations of both center and surround and black and white (high-contrast) stimulus, the estimated optimal parameters allow the freezing percept until a center velocity of about 20°/s, which is still higher than what Wertheim and Paffen found in their experiments but lower compared to our sinusoidal stimulation with the airplane and the greenhouse. The fact that the velocities' ratio between surround and center had to be adjusted to obtain the freezing suggests that it might be influenced by several parameters, such as contrast, spatial frequency's content, configuration, and possibly color of the stimulus.

Table IV.70-2. Critical parameters for the freezing rotation illusion. υ_{center} = center's constant angular velocity (°/sec); $v_{\text{Avg Surround}}$ = surround's average angular velocity adjusted (°/sec) inferred from the optimal surround's maximal angular velocity (formula *v*); $\upsilon_{\text{Max Surround}}$ = surround's maximal angular velocity (°/sec); f_{Surround} = surround's angular frequency (Hz); T/2 = duration (sec) the illusion lasts best for constant velocities stimulation. (a) The "ideal" parameters are given for the surround's sinusoidal stimulation according to the given formulas and the display with the plane and the greenhouse. (b) From a rough fit of Wertheim and Paffen (2009) data, we extrapolated the "ideal" surround-center ratio k = 5.27 to get the FRI with the array of circles for the center and the sinusoids for the background. We extrapolated the presentation time from the period of the optimal sinusoidal frequency (i.e., T/2).

Table IV.70-2a.

υ_{Center}	$\upsilon_{\text{Avg Surround}}$	$\upsilon_{\text{Max Surround}}$	f_{Surround}
6	15.82	24.85	0.13
10	23.82	37.42	0.20
12	27.82	43.7	0.23
20	43.82	68.83	0.36
24	51.82	81.4	0.43
30	63.82	100.25	0.53
40	83.82	131.66	0.69
50	103.82	163.08	0.85
60	123.82	194.5	1.02
70	143.82	225.91	1.18

Table IV.70-2b.

υ_{Center}	$\upsilon_{\text{Avg Surround}}$	T/2
2	14.36	4.23
5	30.17	2.01
10	56.52	1.07
12	67.06	0.91
15	82.87	0.73
20	109.22	0.56

When we depart from the optimal ratio k, the central texture is perceived as either rotating in the same direction as its physical rotation for ratios inferior to the optimal one or rotating in the opposite direction for ratios superior to the optimal. However, it is difficult from the current application to make good estimates because of the perceived

increased velocity observed during counterrotations of the center and the background, which might interfere in estimating the center velocity. The only "sizeable" percept occurs when the surround exerts a *capturing effect* on the center (i.e., when the latter seems to have the same velocity as the former, while it actually rotates faster). This occurs for surround-center velocities' ratios inferior to the optimal one. In the given example, the perceived center velocity was as low as 1.67 its physical velocity (47°/s) but stable in the percept (i.e., a velocity equal to that of the surround) over a range of 15°/s velocity changes. This observation goes against Wertheim and Paffen's (2009) claim that the perceived center velocity is underestimated by a fixed amount, probably because the relevant ratios of background-center velocities were not investigated.

For constant rotations, it is also likely that the stimulation period must be adjusted according to the surround angular velocity, so that large displacements provoked by fast velocity are compensated by short display periods and vice versa (Table IV.70-1b). This indicates that the mechanism involved in obtaining a motionless percept measures proportional angular displacement between surround and center rather than their difference as proposed by Wertheim and Paffen (2009).

Mechanisms at Work

In a system parsing objects using motion cues it certainly makes sense to assign three different states to objects (for a recent discussion about object recognition, see DiCarlo & Cox, 2007): (a) objects that do not move, (b) objects that move together with their moving surround, and (c) objects that move on their own. We can consider, at least for rotation, that (a) objects can be perceived as *stationary relative to space* (e.g., FRI) when they rotate in the same direction and at a given velocity much lower than that of the surround; (b) objects can be perceived as *stationary relative to their surround* (i.e., "captured" by the surround) when they rotate in the same direction and at some higher velocities than that of the surround; and (c) objects are classified as *self-displacing*, either when they move in the opposite direction to that of the surround or in the same direction but at lower velocities than the "ideal" one, or when they are stationary but displayed on a moving surround. What are the mechanisms that can give an account for these phenomena?

Similarly to induced motion, perceived motion of the center in our illusions is modified by the surround's velocity but not the opposite, which obeys the principle of enclosure necessary to induced motion proposed by Duncker (Day, 1981; Duncker, 1929). However, we show that the center, when it has an elongated shape, does not need to be within the "surround" area but could be much larger. As long as the central object allows enough of the surround's inducing area to be visible, as with a semitransparent central circle occupying the entire surround, the illusions occur. If it is likely that induced rotational motion would increase or decrease the perceived center velocity depending on whether the surround counter- or corotates, it cannot justify on its own the freezing effect, especially at high velocities, where its impact should be reduced (see, for magnitudes, Day, 1981; Duncker, 1929; Wertheim & Paffen, 2009).

Similar considerations could be made for induced motion aftereffects, in particular when using constant velocity rotations of the background (Anstis & Reinhardt-Rutland, 1976). However, using an oscillating rather than a continuously rotating surround should suppress motion aftereffects because of the masking influence of back-and-forth rotations.

Depending on particular interactions between eccentricity, size, and luminance contrast of the central object, induced motion can change to motion capture; that is, the direction of the illusions is reversed (Murakami & Shimojo, 1993, 1996). Our illusion, where the center appears to stick to the surround, may seem reminiscent of motion capture, but it differs in several respects. Our illusory motion capture was not obtained by modifying the parameters mentioned previously. In particular, it occurs regardless of the object's size and eccentricity, parameters that normally promote motion capture phenomena (Murakami & Shimojo, 1993, 1996; Ramachandran & Cavanagh, 1987; Watanabe & Cavanagh, 1992). Rather it is the velocities' ratio, which seems the crucial determinant to elicit rotation capture. Our visual system seems to require a minimal angular velocity safety margin; before that corotating center and surround are perceived as moving at different velocities. In our example rotation capture appears for ratios where the center's velocity is about 1.2 to 1.7 times the average surround's velocity. Yet with peripheral fixation, the center appears to stick to the surround at larger ranges of velocities' ratios than with central fixation.

The underlying physiological correlates, which can give an account for the FRI and its variants, might find their origin in higher cortical areas than generally reported for induced motion. Neurons, which are able to detect rotation, have been recorded in the macaque's MSTd, the dorsal portion of the middle superior-temporal gyrus (Duffy & Wurtz, 1991a, 1991b; Graziano, Andersen, & Snowden, 1994; Saito et al., 1986; Tanaka & Saito, 1989). If shape, size, and contrast seem irrelevant to influence the magnitude of the responses of these neurons, different cells respond to different ranges of velocities (Tanaka & Saito, 1989). According to functional magnetic resonance imaging studies in humans, the equivalent area of MSTd is located in the ascending limb of the infero-temporal sulcus, known as the MT+ complex (Dukelow et al., 2001). Since the first detectors for rotations are found only at higher levels of the visual system, the mechanisms involved in detecting rotation and other complex motions are likely to be different from those involved in translational motion (Lee & Lu, 2010; Morrone, Burr, & Vaina, 1995). However, to detect rotation, it may suffice to detect orientation changes over time (Benton, O'Brien, & Curran, 2007). In their fractal rotational stimulus, "the rotational characteristics are the same at every point of the image" (Benton, O'Brien, & Curran, 2007, p306); that is, the location of the rotation center is undetermined. However, when looking at the stimulus through an arbitrary aperture, local rotation is perceived with its rotation axis apparent at the middle of the aperture. Here the perception of rotation depends clearly on prior analysis of the borders of the rotating

pattern somewhere downstream of MSTd. The MT is the prime candidate to map the borders of a moving or even rotating object, since it contains cells with center-surround antagonistic receptive fields with respect to motion direction (Born & Tootell, 1992; Murakami & Shimojo, 1996). However, these MT cells project to MSTl (the lateral ventral region of the MST), where cells also have antagonistic receptive fields, rather than to MSTd (Berezovskii & Born, 2000; Born & Tootell, 1992).

CONCLUSIONS

Unfortunately, experimental psychometrical data about the FRI and its variants are scarce, and we can only speculate about the underlying mechanisms. However, neither rotational induced motion nor its aftereffect or motion capture seem to fully explain the illusions we present here. Rather, interactions between two branches of our motion visual system—one concerned with object-surround segregation located in the inferior temporal cortex, which receives inputs from MT and possibly from MSTl "motion border cells," and one concerned with complex (rotational) motion situated in MSTd or higher up—may be a more suitable explanation for our illusions.

REFERENCES

Anstis, S. M., & Reinhardt-Rutland, A. H. (1976). Interactions between motion aftereffects and induced movement. *Vision Research*, *16*(12), 1391–1394.

Becklen, R., & Wallach, H. (1985). How does speed change affect induced motion? *Attention, Perception, & Psychophysics*, *37*(3), 231–236.

Benton, C. P., O'Brien, J. M., & Curran, W. (2007). Fractal rotation isolates mechanisms for form-dependent motion in human vision. *Biology Letters*, *3*(3), 306–308.

Berezovskii, V. K., & Born, R. T. (2000). Specificity of projections from wide-field and local motion-processing regions within the middle temporal visual area of the owl monkey. *The Journal of Neuroscience*, *20*(3), 1157–1169.

Born, R. T., & Tootell, R. B. H. (1992). Segregation of global and local motion processing in primate middle temporal visual area. *Nature*, *357*(6378), 497–499.

Day, R. (1981). Induced rotation with concentric patterns. *Perception & Psychophysics*, *29*(5), 493–499.

DiCarlo, J. J., & Cox, D. D. (2007). Untangling invariant object recognition. *Trends in Cognitive Sciences*, *11*(8), 333–341.

Duffy, C. J., & Wurtz, R. H. (1991a). Sensitivity of MST neurons to optic flow stimuli: I. A continuum of response selectivity to large-field stimuli. *Journal of Neurophysiology*, *65*(6), 1329–1345.

Duffy, C. J., & Wurtz, R. H. (1991b). Sensitivity of MST neurons to optic flow stimuli: II. Mechanisms of response selectivity revealed by small-field stimuli. *Journal of Neurophysiology*, *65*(6), 1346–1359.

Dukelow, S. P., DeSouza, J. F., Culham, J. C., van den Berg, A. V., Menon, R. S., & Vilis, T. (2001). Distinguishing subregions of the human MT+ complex using visual fields and pursuit eye movements. *Journal of Neurophysiology*, *86*(4), 1991–2000.

Duncker, K. (1929). Über induzierte Bewegung (Ein Beitrag zur Theorie optisch wahrgenommener Bewegung). *Psychologische Forschung*, *12*, 180–259.

Dürsteler, M. R. (2005). Eine neuartige Bewegungstäuschung: Duncker illusion und "motion capture" im widerstreit. *Ophta*, *3*, 19–26.

Dürsteler, M. R. (2007). Measuring the freezing rotation illusion. *Journal of Vision*, *7*(9), 986.

Dürsteler, M. R. (2008). The freezing rotation illusion. In C. Kennard & J. R. Leigh (Eds.), *Using eye movements as an experimental probe of brain function* (Vol. 171, pp. 283–285). Amsterdam: Elsevier.

Graziano, M., Andersen, R., & Snowden, R. (1994). Tuning of MST neurons to spiral motions. *The Journal of Neuroscience*, *14*(1), 54–67.

Ito, H. (2012). Illusory object motion in the centre of a radial pattern: The pursuit-pursuing illusion. *i-Perception*, *3*(1), 59–87.

Lee, A. L. F., & Lu, H. (2010). A comparison of global motion perception using a multiple-aperture stimulus. *Journal of Vision*, *10*(4), 9, 1–16.

Morrone, M. C., Burr, D. C., & Vaina, L. M. (1995). Two stages of visual processing for radial and circular motion. *Nature*, *376*(6540), 507–509.

Murakami, I., & Shimojo, S. (1993). Motion capture changes to induced motion at higher luminance contrasts, smaller eccentricities, and larger inducer sizes, *Vision Research*, *33*(15), 2091–2107

Murakami, I., & Shimojo, S. (1996). Assimilation-type and contrast-type bias of motion induced by the surround in a random-dot display: Evidence for center-surround antagonism. *Vision Research*, *36*(22), 3629–3639.

Ramachandran, V. S., & Cavanagh, P. (1987). Motion capture anisotropy. *Vision Research*, *27*(1), 97–106.

Reinhardt-Rutland, A. H. (1981). Peripheral movement, induced movement, and aftereffects from induced movement. *Perception*, *10*(2), 173–182.

Reinhardt-Rutland, A. H. (1988). Induced movement in the visual modality: An overview [Review]. *Psychological Bulletin*, *103*(1), 57–71.

Saito, H., Yukie, M., Tanaka, K., Hikosaka, K., Fukada, Y., & Iwai, E. (1986). Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *The Journal of Neuroscience*, *6*(1), 145–157.

Tanaka, K., & Saito, H. (1989). Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *Journal of Neurophysiology*, *62*(3), 626–641.

Wallach, H., & Becklen, R. (1983). An effect of speed on induced motion. *Attention, Perception, & Psychophysics*, *34*(3), 237–242.

Watanabe, T., & Cavanagh, P. (1992). Depth capture and transparency of regions bounded by illusory and chromatic contours. *Vision Research*, *32*(3), 527–532.

Wertheim, A. H., & Paffen, C. L. E. (2009). Centre-surround relative motion and the freezing rotation illusion. *Perception*, *38*(11), 1610–1620.

Wertheim, A. H., & Reymond, G. (2007). Neural noise distorts perceived motion: The special case of the freezing illusion and the Pavard and Berthoz effect. *Experimental Brain Research*, *180*(3), 569–576.

Chapter 71
Second-Order Reversed Phi

Zhong-Lin Lu and George Sperling

A phi stimulus is composed of a vertical white bar at one location followed after a brief interval by an adjacent white bar. Observers perceive motion from the first to the second bar. The perceived motion in the phi stimulus was famously important to the Gestalt psychologists (e.g., Wertheimer, 1912) as an example of pure motion. When Reichardt (1961) showed a phi stimulus to a beetle (chlorophanus), the beetle also behaved as if it has seen motion from the first to the second bar. However, when the second bar was black instead of white, the beetle responded as if the direction of motion had been reversed. This observation formed the cornerstone of Reichardt's correlation model of the beetle's motion computation, a model that with suitable elaborations also describes human motion computations (van Santen & Sperling, 1984, 1985).

When almost any object is depicted at one location at time 1 (frame 1), and that same object is shown at an adjacent location after a suitable delay (frame 2), it generally produces a vivid sensation of motion. In what has come to be known as a first-order reversed-phi motion stimulus, the black–white contrast of successive frames is reversed; that is, the even frames are the negatives of the odd frames and vice versa. Under suitable conditions, the direction of apparent motion will appear to be reversed (Anstis, 1970; Anstis & Rogers, 1975; Chubb & Sperling, 1989).

In second-order reversed-phi stimuli (Lu & Sperling, 1999; Nishida, 1993), what is reversed between frame 1 and frame 2 is not the black–white contrast of individual pixels but the overall amount of texture-contrast of areas of the stimuli. The stimuli (Fig. IV.71-1) are constructed such that the first-order motion system (the one originally described by Reichardt, 1961) cannot perceive any systematic motion; that is, the stimuli are invisible to first-order motion analysis. In second-order reverse phi, the salient features (areas with a particular texture) that across frames reverse the sign of their difference with the background (e.g., areas that have more texture contrast than the background in frame 1) have less texture contrast than the background in frame 2. Recall that the first-order motion energy model utilizes the contrast of individual pixels including the sign of the contrast, the second-order motion energy model utilizes the total quantity of texture contrast of textured areas without the sign of the contrast, and the third-order motion energy model utilizes the salience of areas (i.e., high-salience features versus low-salience background; Lu & Sperling, 1995a, 1995b, 2001). In the stimuli of Figure IV.71-1, the salient stimulus *features* move in the forward (feature displacement) direction, but the second-order motion energy model predicts motion in the reversed direction.

In all the second-order reversed-phi stimuli shown in Figure IV.71-1, a neutral background texture occupies 75% of the display area. The remaining 25% of the stimulus is filled with a textured rectangle that varies from frame to frame in texture contrast (Fig. IV.71-1a), in texture spatial frequency (Fig. IV.71-1b), or in temporal frequency (i.e., flicker rate; Fig. IV.71-1c).

In Figure IV.71-1a the translating bars are composed of either high-contrast texture or low-contrast texture on a medium-contrast background. In successive frames, the texture bars move one bar width to the right and the high-contrast regions switch to low contrast and vice versa. For the second-order system, increasing and decreasing texture contrast is equivalent to increasing and decreasing luminance in the first-order system (Chubb & Sperling, 1988; Chubb, Sperling, & Solomon, 1989). The stimulus of Figure IV.71-1a is similar to the stimulus of Nishida (1993). Nishida's stimulus involved just two levels of texture contrast; the areas to which these contrast levels were assigned alternated in consecutive frames. The stimulus in Figure IV.71-1a involves three levels of texture contrast, with the middle level remaining the same from frame to frame and the areas of high and low contrast reversing on consecutive frames.

In Figure IV.71-1b the bars are composed of either low spatial frequency sine-wave gratings or high spatial frequency sine-wave gratings on a medium spatial frequency sine-wave grating background. The bars shift one bar width to the right and alternate between high and low spatial frequency textures in successive frames, always with a new random spatial phase. In second-order motion, decreasing spatial frequency is equivalent to increasing texture contrast (Werkhoven, Sperling, & Chubb, 1993, 1994),

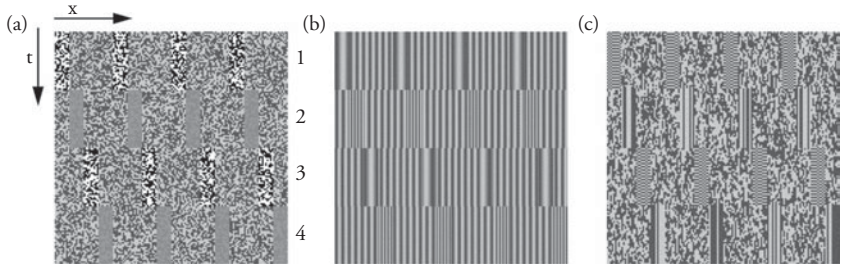


Figure IV.71-1. Three types of second-order reversed-phi stimuli. Each panel shows four consecutive frames (labeled 1, 2, 3, 4). A small rectangular area that occupies 25% of the area is the "figure"; the rest is background. When reproduced properly, figure and background have the same average contrast in all panels; that is, from a sufficiently large viewing distance, the panels look uniformly gray. In each successive frame, the figure translates 90° (one-quarter of the period, the repetition cycle) to the right. For all stimuli, if the texture within the figure produces more overall visual response than the background in one frame, it produces less than the background in the next frame and so on. Panels (a) and (b) illustrate the actual appearance of the stimuli; panel (c) is more schematic. Stimulus (c) consist of vertical lines, somewhat like (b), except that the lines change their contrast from black to white according to the color in the time axis, *t*. (Adopted from Figure 6 in Lu, Z.-L., & Sperling, G. (1999). Second-order reversed phi. *Perception & Psychophysics*, *61*(6), 1075–1088.)

so these stimuli, too, become reversed-phi stimuli for the second-order system.

Figure IV.71-1c represents flickering stimuli. The vertical axis represents time. All the stimuli are composed entirely of vertical stripes; the *x, t* graph represents the color of a stripe at a particular instant in time. For the first- and third-quarter cycles of the stimulus (rows 1 and 3), the bars represent regions with a maximum rate of pixel flicker (a black–white reversal of every pixel on every frame), superimposed on a background that has a random, medium rate of flicker. In the second- and fourth-quarter cycles, bars remain unchanged during their entire exposure (vertical stripes in panel c, low flicker rate regions). The translating bars displace one bar width to the right, switching from high to low flicker and vice versa from frame to frame.

The second-order reversed phi stimuli of Figure IV.71-1 were shown to subjects at rates between 0.94 and 15 Hz (Lu & Sperling, 1999). In peripheral vision, every observer saw presentations only in the reversed direction (opposite to the bar displacement). In some conditions—for example, at the highest temporal frequencies of the flicker modulation stimuli—there were some trials in which no motion was perceived. When motion was perceived, however, it was universally in the direction of second-order reversed phi.

In central viewing, observers reported perceiving motion in the forward direction at low temporal frequencies and in the reversed direction at high temporal frequencies. The temporal frequency at which the direction of perceived motion changed depended on the stimulus but was similar for all observers. For the contrast modulation stimulus (Fig. IV.71-1a), all observers perceived motion in the reversed direction at 15 Hz and motion in the forward direction at temporal frequencies of 0.94, 1.88, and 3.75 Hz; 7.5 Hz was ambiguous for all observers. For the spatial frequency modulation stimulus (Fig. IV.71-1b), all observers reported forward motion at temporal frequencies of 0.94 and 1.88 Hz and reversed motion at 3.75 and 7.50 Hz. For the flicker modulation stimulus (Fig. IV.71-1c), at temporal frequencies of 0.94 and 1.88 Hz, all observers reported perceiving motion in the forward direction. At the highest resolvable temporal frequency for these stimuli (3.75 Hz), all observers reported perceiving motion in the reversed direction. Moving the observer away from the displays had the same effect as changing from central viewing to periphery viewing: reversed motion became more dominant.

To understand second-order reversed phi, we start with a theory on first-order reversed phi. Figure IV.71-2a shows four frames of a first-order reversed-phi stimulus (Chubb & Sperling, 1989) in which the contrast of the bars alternates polarity (from white to black and from black to white) from frame to frame. Applying motion energy analysis directly to raw stimulus contrast predicts a motion direction that is opposite to the bar displacement (reversed phi) because the dominant Fourier component (indicated by the light and the dark stripes of a superimposed sine-wave grating) is oriented in the leftward direction (Fig. IV.71-2b). When viewing in the periphery or from afar, human observers perceive this so-called reversed motion direction (as predicted by motion energy analysis). In central vision, humans perceive motion in the forward direction, contradicting what is predicted by motion energy analysis. This was explained by a model in which stimulus contrast is first full-wave rectified (absolute value computation) and only then subjected to motion energy analysis (Fig. IV.71-3). After rectification, the dominant Fourier component is in the forward direction (indicated by the superimposed sine-wave grating, Fig. IV.71-2c).

A second-order motion energy model consists of a texture grabber followed by a motion energy detector (Fig. IV.71-3). The texture grabber consists of a spatiotemporal linear filter. The spatial component is represented as a cosine receptive field (which is primarily sensitive to a particular orientation and spatial frequency), and the temporal component is a bandpass filter that is relatively insensitive to very low and very high temporal frequencies. Full-wave rectification (e.g., absolute value or square law rectifier; absolute value is shown) gives the texture grabber an output that depends only on the total quantity of texture in its preferred orientation and frequency, independent of the sign of the outputs of the linear filters, which may be either positive or negative. Motion energy analysis determines the direction of motion of the texture detected by the texture grabber.

The second-order reversed phi stimuli in Figure IV.71-1 are preprocessed by a texture grabber that consists of a low-pass spatial filter and a band-pass temporal filter followed by full-wave rectification (Fig. IV.71-3). The texture grabber

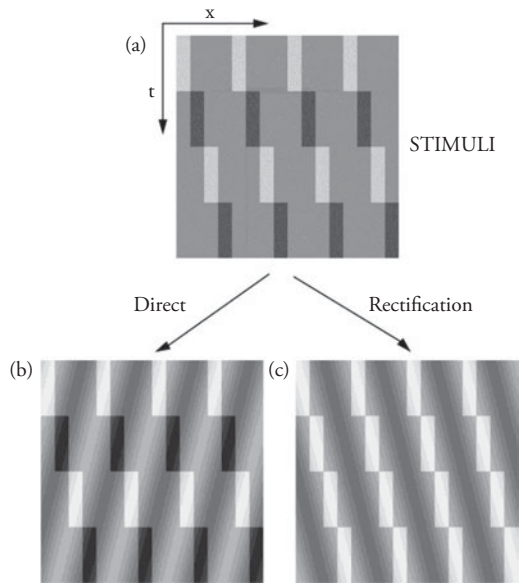


Figure IV.71-2. First-order reversed-phi: stimulus and theory. The horizontal axis represents space; the vertical axis represents time. (a) A reversed-phi square-wave stimulus. (b) The fundamental sine wave from a Fourier analysis of the stimulus (which is the primary stimulus to the motion system) is shown superimposed on the stimulus, illustrating that it signals a direction opposite to the direction in which figure moves. (c) The second-order motion system is indifferent to the sign of contrast, so all stimulus contrasts are shown as equal and positive. The third-order system is sensitive to the figure, so all areas of the figure are shown as positive. The superimposed fundamental sine wave illustrates both higher-order motion systems' report of motion in the direction of the figure. (After Lu, Z.-L., & Sperling, G. (1999). Second-order reversed phi. *Perception & Psychophysics*, *61*(6), 1075–1088, with permission.)

produces an output that is equivalent to a first-order reversed-phi stimulus (Fig. IV.71-4). Rectification alone would have sufficed to reveal the second-order motion of the texture contrast modulation stimulus in Figure IV.71-1a. However, rectification must be preceded by low-pass spatial filtering to expose the motion of the three grating frequencies stimulus of Figure IV.71-1b. Temporal band-pass or temporal low-pass filtering followed by rectification (Chubb & Sperling, 1988) is required to expose the motion of the flicker stimulus of Figure IV.71-1c. Figure IV.71-4 illustrates that, after preprocessing, the dominant motion energies of the reversed-phi stimuli of Figures IV.71-1a, IV.71-1b, and IV.71-1c are in the leftward direction, opposite to the rightward displacement of the bars. The motion that is extracted from second-order reversed-phi stimuli by a texture grabber, followed by a motion energy computation, would be reversed motion.

Because the second-order reversed-phi stimuli are invisible to first-order analysis, and because second-order analysis produces motion in the reversed direction, the perception of forward motion requires yet another mechanism, presumably the third-order salience map mechanism proposed by Lu and Sperling (1995a, 1995b, 2001). The three reversed-phi stimuli in our experiments were constructed so that one-fourth of the area was occupied by the figure, which changed type from frame to frame, and three-fourths of the area was occupied by the background, which remained physically unchanged from frame to frame in the contrast modulation stimuli and unchanged in type for the spatial frequency and flicker stimuli. The third-order motion mechanism is assumed by Lu and Sperling (1995a, 1995b) to compute the moment-to-moment motion of those parts of the visual stimulus that are marked as figure in a neural salience field. So if (a) occupying a small part of the field and changing from frame to frame were to

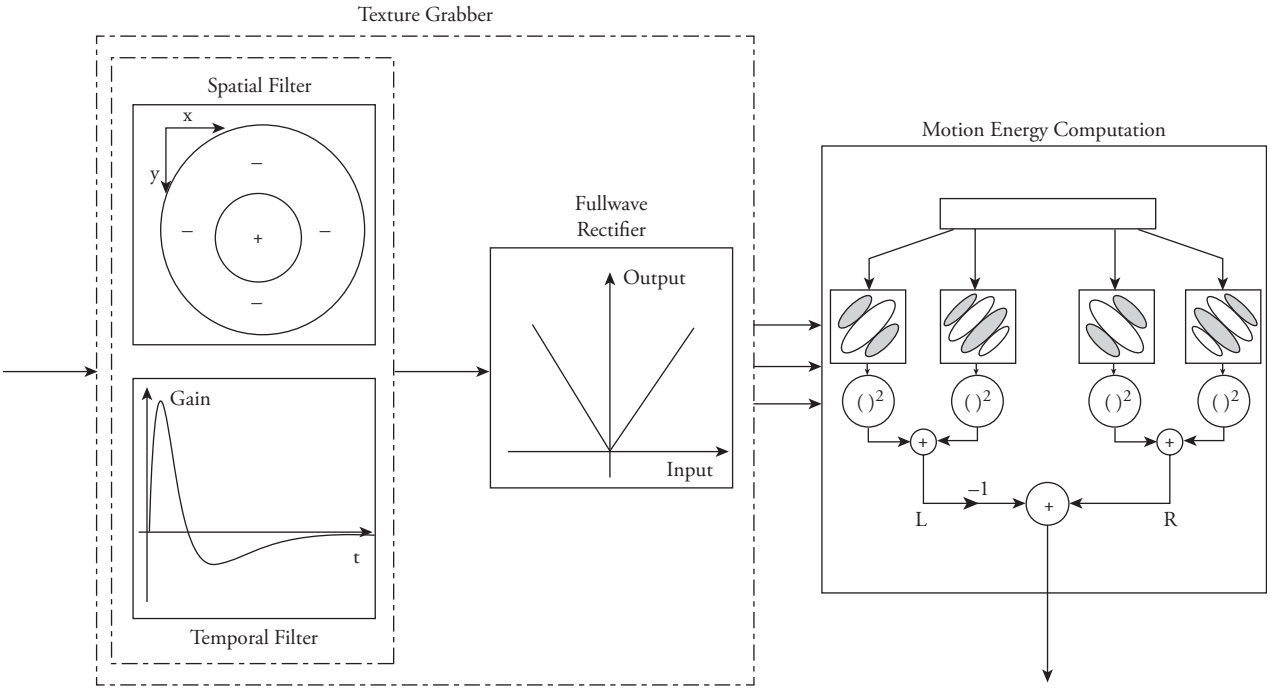


Figure IV.71-3. A second-order motion energy model. A texture grabber is followed by a motion energy detector. A texture grabber consists of a two-dimensional bandpass spatial filter and a one-dimensional bandpass temporal filter followed by rectification. The motion energy computation is tuned to optimally compute motion power in a three-dimensional spatiotemporal frequency band. (After Lu, Z.-L., & Sperling, G. (1999). Second-order reversed phi. *Perception & Psychophysics*, *61*(6), 1075–1088, with permission.)

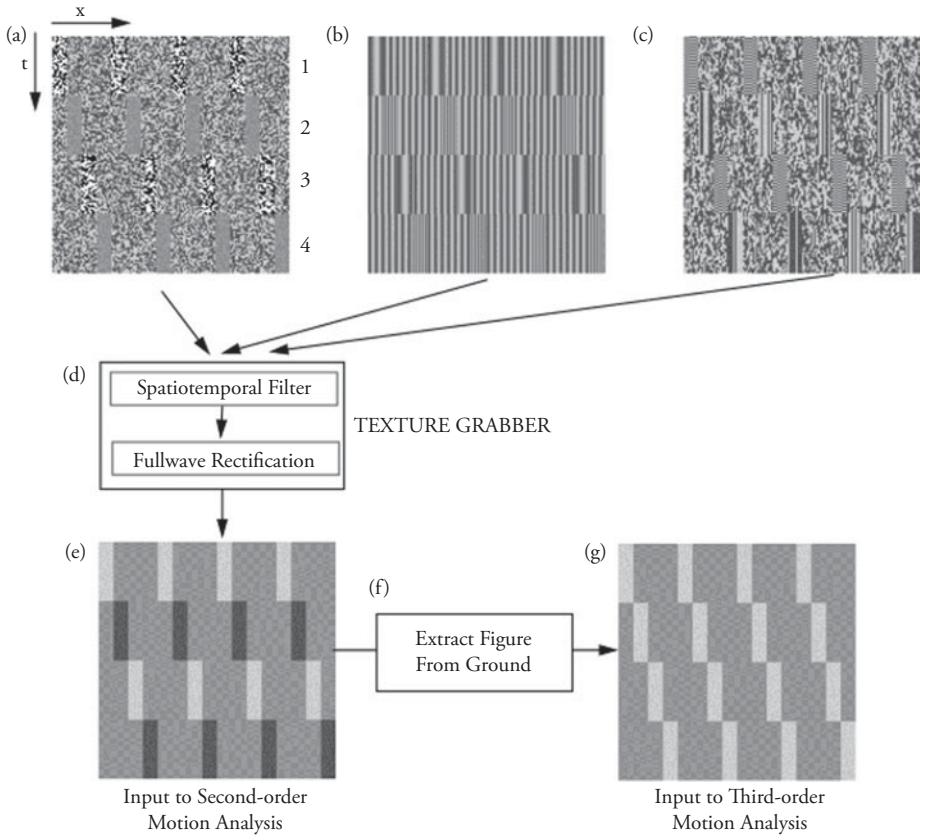


Figure IV.71-4. (a) through (c) Second-order reversed phi stimuli. (d) Spatiotemporal filtering and full-wave rectification (see Fig. IV.71-3). (e) The stimuli, as transformed by (d), can be represented as a space–time graph. For stimuli (a), (b), and (c), (e) can also be interpreted as representing four consecutive frames of texture grabber output. The dominant motion energy in panel (e) is leftward (opposite to the bar displacement direction). Applying motion energy analysis on the output of the texture grabber would predict motion in the reversed direction. (f) Salience computation. (g) A salience system that marked the location of features that differed from the background in each frame would compute motion in the forward (feature) direction. (After Lu, Z.-L., & Sperling, G. (1999). Second-order reversed phi. *Perception & Psychophysics*, *61*(6), 1075–1088, with permission.)

cause that portion of the stimulus to be marked as figure, and (b) occupying a large part of the field and remaining stable were to cause that portion of the stimulus to be marked as ground, and (c) the third-order motion system computed the motion of figure, then the second-order reversed-phi stimuli would support third-order motion in the forward direction. This is precisely what occurred.

The third-order motion system has a much lower cutoff temporal frequency (≈ 3–4 Hz) than does the second-order motion system (≈ 10–12 Hz; Lu & Sperling, 1995a). Similarly, the higher the order of the motion system, the lower its spatial resolution and, therefore, the more likely it will be confined to areas near the fovea, where spatial frequency resolution is better. These results—that the third-order motion direction was (a) never perceived in peripheral viewing and (b) perceived in foveal areas only at low temporal frequencies—are entirely consistent with the previously noted properties of the second- and third-order motion systems.

REFERENCES

Anstis, S. (1970). Phi movement as a subtraction process. *Vision Research*, *10*(12), 1411–1415.

Anstis, S. M., & Rogers, B. J. (1975). Illusory reversal of visual depth and movement during changes of contrast. *Vision Research*, *15*(8), 957–956.

Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli- A general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America A*, *5*, 1986–2007.

Chubb, C., & Sperling, G. (1989). Two motion perception mechanisms revealed through distance-driven reversal of apparent motion. *Proceedings of the National Academy of Sciences USA*, *86*(8), 2985–2989.

Chubb, C., Sperling, G., & Solomon, J. A. (1989). Texture interactions determine perceived contrast. *Proceedings of the National Academy of Sciences USA*, *86*(23), 9631–9635.

Lu, Z.-L., & Sperling, G. (1995a). Attention-generated apparent motion. *Nature*, *377*(6546), 237–239.

Lu, Z.-L., & Sperling, G. (1995b). The functional architecture of human visual motion perception. *Vision Research*, *35*(19), 2697–2722.

Lu, Z.-L., & Sperling, G. (1999). Second-order reversed phi. *Perception & Psychophysics*, *61*(6), 1075–1088.

Lu, Z.-L., & Sperling, G. (2001). Three-systems theory of human visual motion perception: Review and update. *Journal of the Optical Society of America A*, *18*(9), 2331–2370.

Nishida, S. Y. (1993). Spatiotemporal properties of motion perception for random-check contrast modulations. *Vision Research*, *33*(5), 633–645.

Reichardt, W. (1961). Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In W. A. Rosenblith (Ed.), *Sensory communication* (pp. 303–317). Cambridge, MA: MIT Press.

van Santen, J. P., & Sperling, G. (1984). Temporal covariance model of human motion perception. *Journal of the Optical Society of America A*, *1*(5), 451–473.

van Santen, J. P., & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A*, *2*, 300–321.

Werkhoven, P., Sperling, G., & Chubb, C. (1993). The dimensionality of texture-defined motion: A single channel theory. *Vision Research*, *33*(4), 463–485.

Werkhoven, P., Sperling, G., & Chubb, C. (1994). Perception of apparent motion between dissimilar gratings: Spatiotemporal properties. *Vision Research*, *34*(20), 2741–2759.

Wertheimer, M. (1912). Über das Denken von Naturvölker: 1. Zahlen und Zahlengebilde. *Zeitschrift fuer Psychologie und Physiologie der Sinnesorgane*, *60*, 321–389.

Chapter 72
Attention-Generated Apparent Motion

Zhong-Lin Lu and George Sperling

In attention-generated apparent motion (AGM), a display that appears to flicker but does not appear to have any obvious direction of motion will suddenly appear to move in a consistent direction when a particular feature in the display is attended. In the AGM displays illustrated in this chapter (Figs. IV.72-1, IV.72-2, and IV.72-3), attention causes apparent motion in either one of two opposite directions depending on the feature attended, such as left versus right slanting lines, black versus white dots, or red versus green stripes (Blaser, Sperling, & Lu, 1999; Lu & Sperling, 1995a). A variant of the procedure involves attending to a feature in a different task for an extended time period before the motion task. For example, looking for a red target in a visual search task for about an hour of total viewing time caused the direction of perceived motion in subsequently viewed ambiguous motion displays to remain biased in the "red" direction for a month (Tseng, Gobell, & Sperling, 2004).

The AGM illusion was created to demonstrate the existence of and reveal the mechanisms of feature-salience motion (third-order motion) perception. This required stimuli that were invisible to first- and second-order motion processes. The critical aspect was that only a feature's salience was marked from frame to frame, not the feature itself. A motion system transmits information only about the location, direction, and speed of movement. Information about what is moving—the features themselves—is carried by a pattern processing system. The key to demonstrating feature-salience motion is to use voluntary attention to influence which features are marked as salient in alternating-feature displays, in which frames made of one type of stimulus material are interleaved with frames made of another, entirely different type of material (Cavanagh, Arguin, & von Grunau, 1989). It does not matter if the actual features being marked change from frame to frame, as long as the marked locations follow a consistent motion trajectory.

The AGM illusion uses alternating feature displays with the following properties. First, motion is invisible to the first-order (Adelson & Bergen, 1985; Reichardt, 1957; Van Santen & Sperling, 1984) and second-order (Cavanagh & Mather, 1989; Chubb & Sperling, 1989) motion perception systems and would be visible only to a feature-salience motion system (Lu & Sperling, 1995a, 1995b, 2001b). Perceiving motion requires combining salience information from the two different materials. Second, one stimulus feature is inherently more salient in one type of frames. Third, the to-be-attended features are entirely ambiguous (have equal salience) in the other type of frames. Fourth, verbal instructions direct the subject to attend only to one of the two kinds of features in the feature-ambiguous frames, and fifth, opposite motion directions are perceived depending on which feature is attended in the equal-salience frames.

Three different kinds of alternating-feature display stimuli are illustrated: depth/texture gratings (Fig. IV.72-1), full-wave/half-wave gratings (Fig. IV.72-2), and full-wave/color gratings (Fig. IV.72-3). In the depth gratings (Figs. IV.72-1b, IV.72-1d), for most observers, the near-appearing peaks are naturally more salient than the background valleys; however, the two textures in Figures IV.72-1a, IV.72-1c, and IV.72-1e have equal salience. In the full-wave gratings (Fig. IV.72-2), the high-contrast areas are naturally more salient than the low-contrast areas. Half-wave gratings (Solomon & Sperling, 1994) are made of alternating areas of white dots on a dark grey background and black dots on a light grey background. After calibration for each subject (Lu & Sperling, 2001a), the areas of white and of black spots have equal salience. In the isoluminant chromatic gratings (Fig. IV.72-3), after isoluminant calibration for each subject (Lu & Sperling, 2001a), the relative salience of the red and green areas depends on the ratio of their chromaticity; the red and green areas can be made to have equal salience by adjusting the ratio (Blaser et al., 1999; Claeys, Lindsey, De Schutter, & Orban, 2003; Lu, Lesmes, & Sperling, 1999).

In all three alternating-feature displays, the phase shift between successive frames is 90°, thereby successive even-numbered frames in the sequence are separated by 180°, as are successive odd-numbered frames. Motion between frames separated by 180° is totally ambiguous. The perception of coherent motion would require combining information from even and odd frames, which are separated by 90°, the ideal separation for motion detection (Van Santen & Sperling, 1984). For each given alternating-feature motion display, the

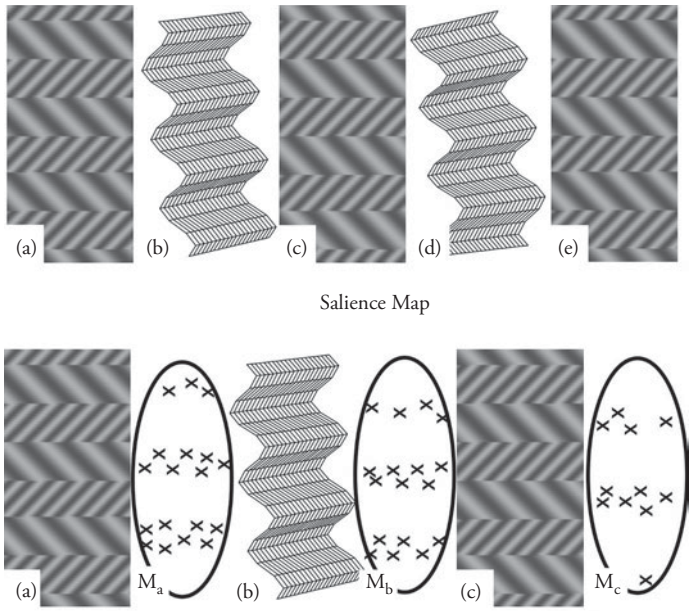


Figure IV.72-1. The depth/texture alternating feature display. First row: A sequence of five superimposed consecutive frames (a–e); the grating pattern in each frame is displaced by 90° from the previous one. The texture frames (a, c, e) are shown as they appeared to the observer; the depth stereograms, (b, d; uniformly textured surfaces that vary in perceived depth) are indicated schematically. Second row: First three frames (a, b, c) and a feature-salience map associated with each frame. The most salient features of the depth frames are the near peaks (upper peaks in the panels) marked by Xs in the salience map (M_b). When a subject attends to the coarse grating, the coarse stripes are marked in the salience map (Xs in M_a and M_c). The direction of perceived motion (downward) follows the space-time trajectory of the Xs as indicated by the dashed line from upper left to lower right. The opposite direction of motion has no support in M_a and M_c; perceiving this bottom-to-top motion would require attending to the fine stripes. (After Lu, Z. L., & Sperling, G. (1995a). Attention-generated apparent motion. *Nature*, *377*(6546), 237–239, with permission.)

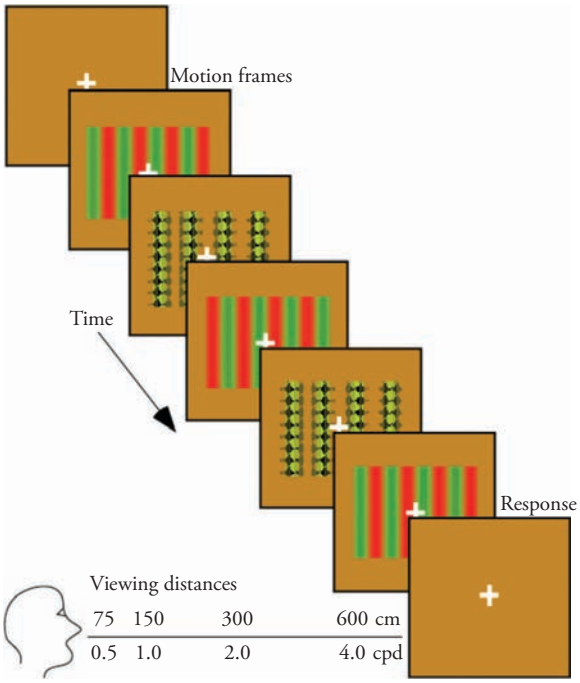


Figure IV.72-3. Five frames of the fullwave/chromatic gratings stimuli. The term "full-wave" merely indicates alternating areas of high- and low-contrast random binary noise; chromatic gratings are isoluminant color gratings. (Sperling, G., Reeves, A., Blaser, E., Lu, Z.-L., & Weichselgartner, E. (2001). Two computational models of attention. In J. Braun, C. Koch, & J. L. Davis (Eds.), Visual attention and cortical circuits. Cambridge, MA: MIT Press. Pp. 177–214, with permission.)

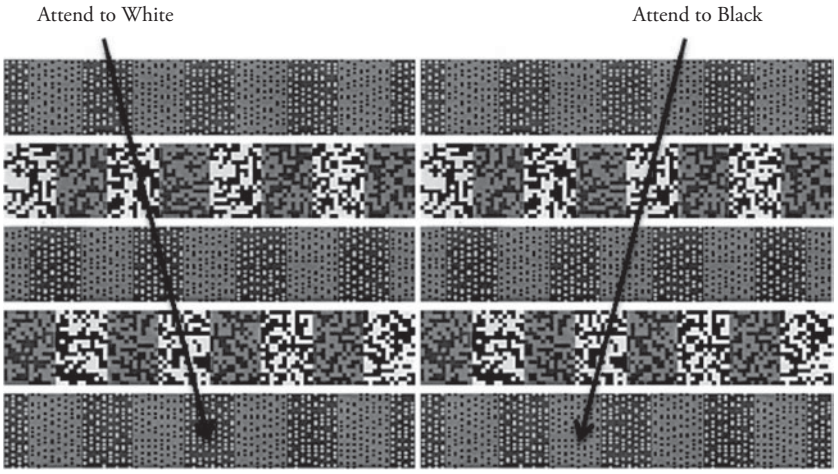


Figure IV.72-2. Left column: Five frames of the full-wave/half-wave stimuli. The term "full-wave" indicates alternating areas of high- and low-contrast random binary noise; "half-wave" indicates alternating areas of white dots on a dark grey background and black dots on a light grey background. Both frames have the same expected luminance everywhere and therefore do not stimulate the first-order (luminance) motion system. Prior to any attention instructions, the half-wave frames are calibrated individually for each subject to find a contrast ratio of white to black for which the direction of perceived motion of the displays was completely random. Right column: The trajectory from high-contrast areas of full-wave frames (which are automatically salient) to white spots of half-wave frames (made salient by attention) produces perceived motion to the right. Within an experimental session, only one type of attention instruction is given. (After Lu, Z. L., & Sperling, G. (2001). Sensitive calibration and measurement procedures based on the amplification principle in motion perception. *Vision Research*, *41*(18), 2355–2374, with permission.)

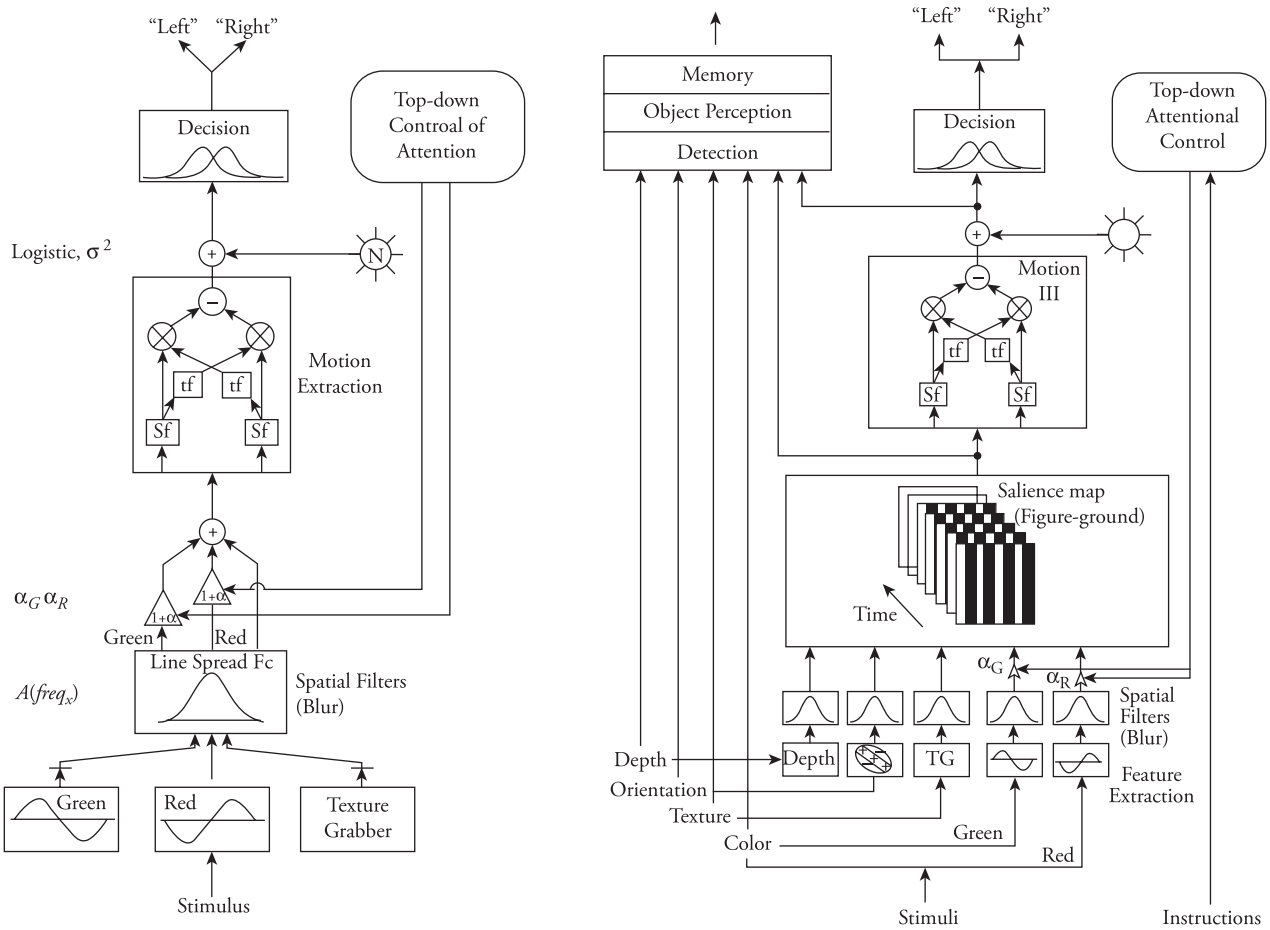


Figure IV.72-4. Models for attention processes in third-order motion. (a) A computational model of attention processes in third-order motion that in (b) is embedded in a more comprehensive model of visual processing. The inputs to the model are stimuli and attention instructions; the computational output is a direction-of-motion judgment. There are two types inputs: visual stimuli and attention instructions. Stimulus inputs are analyzed along various dimensions: depth, orientation, texture, and color channels are indicated. For the present experiments, it is only necessary to consider color (red and green) and texture processing. Instructions to attend to a color are assumed in only the salience pathway to increase the gain of the attended color signal to a value greater than 1.0 so that the attended color input is amplified by $1 + \alpha_G$ or $1 + \alpha_R$. The salience map is the sum of all the stimulus inputs in the salience pathway; its output goes to the Motion III (third-order) computation and (in b) also joins the stimulus inputs in the object-processing pathway. The third-order motion computations is represented as a Reichardt model. It produces a real-valued output that indicates a direction of motion and is perturbed by additive noise *N*. A decision process outputs a response "right" if its input is greater than a criterion and "left" otherwise. (b) A more general model that shows sensory inputs bypassing the salience computation en route to subsequent processing. Although high salience does not seem to perturb the appearance of objects, it does eventually determine what signals are analyzed and remembered. The third-order motion signal is also available to subsequent perceptual processes, as indicated. (After Blaser, E., Sperling, G., & Lu, Z. L. (1999). Measuring the amplification of attention. *Proceedings of the National Academy of Sciences USA*, *96*(20), 11681, and Sperling, G., Reeves, A., Blaser, E., Lu, Z.-L., & Weichselgartner, E. (2001). Two computational models of attention. In J. Braun, C. Koch, & J. L. Davis (Eds.), Visual attention and cortical circuits. Cambridge, MA: MIT Press. Pp. 177–214, with permission.)

perceived motion direction depends on which stimulus feature is being attended in the equal-salient frames.

Lu and Sperling (1995a) found that, at a rate of 67 ms/frame, some subjects immediately perceived motion in a direction consistent with the attention instruction. Others did not. However, after several hundred trials, every subject who participated in the experiment perceived motion in the direction consistent with the attention instruction. At this point, the instruction was reversed (i.e., attend to the previously unattended texture), and trials were continued until performance again stabilized. The task is easier at slower frame rates (e.g., 133 ms/frame). Blaser et al. (1999) showed that the perceived direction from the full-wave/chromatic gratings display depends on both the ratio of chromaticity of the red-green grating and attention instruction; attention amplifies the salience of the attended color. Tseng et al. (2004) showed that, after about one hour of searching for red targets in a search experiment, perceived motion direction from the full-wave/chromatic gratings display remained biased in the red direction for a month.

The latent motion in AGM displays is invisible to first- or second-order motion mechanisms, and the direction of apparent motion depends on the particular feature attended. The illusion suggests the mechanism of third-order motion: At each moment in time, the *x, y* locations of the most significant features are marked (the salience map; Itti, Koch, & Niebur, 1998; Koch & Ullman, 1985). For example, when figure–ground distinctions are meaningful, being marked leads to being labeled "figure." Motion is computed by standard algorithms (Adelson & Bergen, 1985; Reichardt, 1957; Van Santen & Sperling, 1984) from the spatiotemporal changes of the map (Fig. IV.72-4a).

AGM and other related phenomena are all encompassed within the salience map theory (Fig. IV.72-4b). A dynamic salience map of the locations of the most salient stimulus features is determined jointly by stimulus strength (bottom-up) and by selective attention (top-down). Motion is computed directly and automatically from the salience map, but the map also can be used to guide other processes such as object recognition in attention-guided search and memory storage in location-cued recall tasks (Shih & Sperling, 2002; Sperling, 1960).

REFERENCES

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, *2*(2), 284–299.

Blaser, E., Sperling, G., & Lu, Z. L. (1999). Measuring the amplification of attention. *Proceedings of the National Academy of Sciences USA*, *96*(20), 11681.

Cavanagh, P., Arguin, M., & von Grunau, M. (1989). Interattribute apparent motion. *Vision Research*, *29*(9), 1197–1204.

Cavanagh, P., & Mather, G. (1989). Motion: The long and short of it. *Spatial Vision*, *4, 2*(3), 103–129.

Chubb, C., & Sperling, G. (1989). Two motion perception mechanisms revealed through distance-driven reversal of apparent motion. *Proceedings of the National Academy of Sciences USA*, *86*(8), 2985.

Claeys, K. G., Lindsey, D. T., De Schutter, E., & Orban, G. A. (2003). A higher order motion region in human inferior parietal lobule: evidence from fMRI. *Neuron*, *40*(3), 631–642.

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(11), 1254–1259.

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, *4*(4), 219–227.

Lu, Z. L., Lesmes, L. A., & Sperling, G. (1999). Perceptual motion standstill in rapidly moving chromatic displays. *Proceedings of the National Academy of Sciences USA*, *96*(26), 15374.

Lu, Z. L., & Sperling, G. (1995a). Attention-generated apparent motion. *Nature*, *377*(6546), 237–239.

Lu, Z. L., & Sperling, G. (1995b). The functional architecture of human visual motion perception. *Vision Research*, *35*(19), 2697–2722.

Lu, Z. L., & Sperling, G. (2001a). Sensitive calibration and measurement procedures based on the amplification principle in motion perception. *Vision Research*, *41*(18), 2355–2374.

Lu, Z. L., & Sperling, G. (2001b). Three-systems theory of human visual motion perception: review and update. *Journal of the Optical Society of America A*, *18*(9), 2331–2370.

Reichardt, W. (1957). Autokorrelations-auswertung als Funktionsprinzip des Zentralnervensystems. *Zeitschrift Naturforschung Teil A*, *12*, 756.

Shih, S. I., & Sperling, G. (2002). Measuring and modeling the trajectory of visual spatial attention. *Psychological Review*, *109*(2), 260.

Solomon, J. A., & Sperling, G. (1994). Full-wave and half-wave rectification in second-order motion perception. *Vision Research*, *34*(17), 2239–2257.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, *74*(11), 1.

Tseng, C., Gobell, J. L., & Sperling, G. (2004). Long-lasting sensitization to a given colour after visual search. *Nature*, *428*(6983), 657–660.

Van Santen, J. P. H., & Sperling, G. (1984). Temporal covariance model of human motion perception. *Journal of the Optical Society of America A*, *1*(5), 451–473.

Chapter 73
Two-Stroke Apparent Motion

George Mather

THE EFFECT

One hundred years ago, the Gestalt psychologist Max Wertheimer published the first detailed study of the apparent visual movement seen when two adjacent lights are flashed sequentially (see Kolers, 1972). Apparent movement can be created by presenting any two identical pictures successively, with the second shifted slightly in position relative to the first. Figure IV.73-1a shows a very simple example: a bright bar on a dark background is plotted in cross-section, displacing to the right from frame 1 (F1) to frame 2 (F2; solid arrow). The displacement evokes a perception of apparent motion to the right (dashed arrow). Stop-frame animation is based on this well-established technique.

Anstis (1970) reported the surprising observation that if the second picture is the photographic negative of the first, then the apparent direction of movement generated by the two-frame sequence is reversed; rightward spatial displacements are seen as leftward apparent motion and vice versa. Figure IV.73-1b shows a rightward shifting bright bar that reverses contrast in frame 2; apparent motion is now leftward. Anstis called this effect "reversed phi." Anstis and Rogers (1975) found that small luminance edge displacements of about 0.17 arc degree gave the strongest "reversed phi" effect. Anstis and Rogers (1986) later created apparently continuous unidirectional apparent motion by repetitively cycling through four animation frames containing alternation between forward phi and reversed phi sequences. A simple instance of this four-stroke cycle is shown in Figure IV.73-1c. Transitions F1–F2 and F3–F4 produce forward phi, and transitions F2–F3 and F4–F5/1 produce reverse phi. Four-stroke apparent motion sequences generate strong motion aftereffects (MAEs) following prolonged inspection (subsequently viewed stationary patterns appear to move in the opposite direction).

In the 1990s several papers reported another curious visual effect seen in two-frame apparent motion displays: interframe interval (IFI) reversal. When the two frames of the animation are separated by a brief IFI that contains a blank field, the apparent direction of the sequence reverses (Shioiri & Cavanagh, 1990; Strout, Pantle, & Mills, 1994; Takeuchi & De Valois, 1997). So a rightward spatial displacement from the first frame to the second appears as a leftward movement when a brief blank field separates the two frames, as illustrated in Figure IV.73-1d. Reversals of apparent direction only occur over a narrow range of IFIs lasting around 30 to 50 msec. Mather (2006) exploited IFI reversal to create a new illusion of apparently continuous motion called two-stroke apparent motion. The sequence involves only two pattern frames, one spatially displaced relative to the other, which alternate repetitively. Every second pattern frame is followed by a brief blank field (a simple example is shown in Fig. IV.73-1e). Displacements that are not interrupted by the IFI yield apparent motion in the same direction as the spatial displacement (e.g., frame transition F1–F2 in Fig. IV.73-1e). Displacements separated by the IFI yield apparent motion in the direction opposite to the spatial displacement (e.g., frame transition F2–F3/1 in Fig. IV.73-1e, where F3 is identical to F1). The repeating sequence thus appears to move only in one direction (unidirectionally), because forward displacements are seen as forward motion, while backward displacements are apparently reversed and also seen as forward motion. Shifting the IFI so that it follows the first pattern frame rather than the second reverses the direction of the apparent motion seen in the sequence without altering the contents of the two pattern frames themselves. As in the case of four-stroke apparent motion, prolonged exposure to two-stroke apparent motion leads to a strong MAE.

Two-stroke apparent motion is so robust that it can be created relatively easily using any pair of pictures containing spatial displacements, such as two adjacent frames from a video. One of the two frame transitions in the repeating cycle should include an IFI lasting about 30 msec (two display refreshes on a 60 Hz monitor) and containing a blank field set to approximately match the mean luminance of the pattern frames.

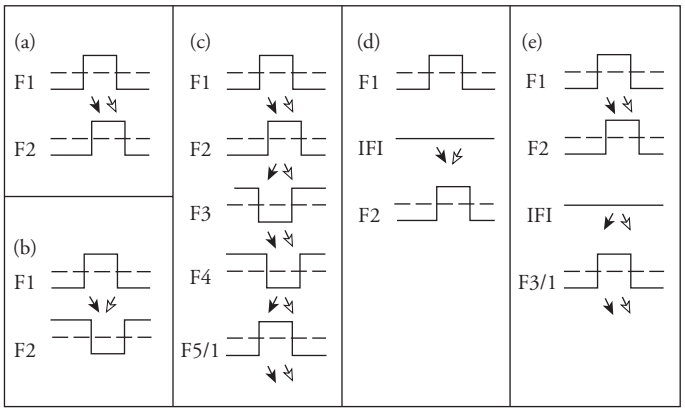


Figure IV.73-1. Frame sequences to create simple motion illusions. (a) When a bright bar displaces to the right (solid arrow) from frame 1 (F1) to frame 2 (F2), apparent motion is seen to the right (outline arrow). (b) When a bright bar displaces to the right and reverses in contrast, apparent motion is seen to the left. (c) A four-stroke cycle in which rightward displacements preserve contrast and leftward displacements reverse contrast. Frame 5 is identical to frame 1 (F5/1) so the sequence repeats, creating an impression of consistent rightward motion. (d) IFI reversal in which two-frame rightward displacement of a bright bar is interrupted by a brief uniform interstimulus interval (IFI). Apparent movement is seen to the left. (e) A two-stroke display containing a bar that displaces to the right from frame 1 to frame 2. Then, after a brief uniform interstimulus interval (IFI), frame 1 reappears (F3/1) and the cycle repeats. The bar appears to move consistently to the right. (Figures by George Mather.)

RELEVANCE OF TWO-STROKE MOTION FOR VISION RESEARCH

On a practical level, the two-stroke motion effect allows one to create dynamic images that appear to move consistently in one direction from the bare minimum of displacement information—namely, just two animation frames. Two-stroke motion is counterintuitive and theoretically challenging: How can the addition of a blank interval to a back-and-forth animation sequence create an impression of continuous forward movement when the blank interval itself contains no useful information? The illusion can be used as a tool to probe the neural computations underlying motion perception. The adequacy of theoretical models of the process can be tested by comparing their output to data on two-stroke motion obtained from psychophysical experiments. The next section summarizes some of the critical stimulus parameters for obtaining two-stroke apparent motion, and the final section considers the detailed implications of the effect for theoretical models.

IMPORTANT PARAMETERS IN TWO-STROKE MOTION

Mather and Challinor (2009) used MAE duration following adaptation to two-stroke motion as an index of the strength of the illusion. They found that the illusion was strongest when IFI duration fell between 30 and 70 msec and when the luminance of the blank IFI matched the mean luminance of the pattern frames. These results mirror those obtained for IFI-reversal motion, indicating a common origin.

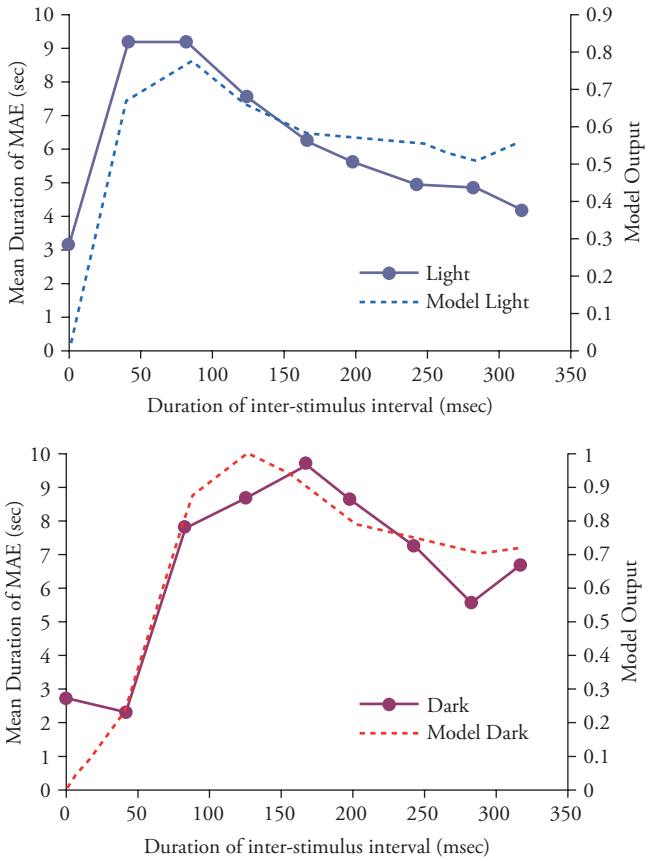


Figure IV.73-2. Solid lines: Results of experiments to measure the duration of the MAE following adaptation to two-stroke motion, as a function of interframe interval (horizontal axis) and mean light level. Data in the top graph were obtained at a mean luminance of 46 cd/m^2; data in the bottom graph were obtained at a mean luminance of 0.23 cd/m^2. Broken lines: Output of a computational model of early motion detection (Adelson & Bergen, 1985), with temporal filter parameters set to the best-fitting values. (Data are replotted from Challinor & Mather, 2010.) (Figures by George Mather.)

Motion processing is known to change significantly at low light levels (e.g., Gegenfurtner, Mayser, & Sharpe, 2000), so if two-stroke motion depends on these processes, it should be sensitive to changes in light level. Challinor and Mather (2010) investigated the effect of varying the mean luminance of the entire pattern. MAEs following exposure to two-stroke motion were measured in bright conditions (the mean luminance of both the pattern frames and the IFI was 46 cd/sq/m.) and in dim conditions (mean luminance 0.23 cd/sq/m). Lower mean luminance appeared to slow down the response of the motion system, stretching out the time dependence of the two-stroke effect in a way that is consistent with other effects of low luminance levels on motion perception: in bright conditions (solid lines in Fig. IV.73-2, top), MAEs were maximal at IFIs of around 60 msec, while in dim conditions (solid lines in Fig. IV.73-2, bottom), the peak in MAE duration occurred at IFIs of around 167 msec.

DISCUSSION OF THEORIES

As Anstis and Rogers (1986) noted, Fourier theory offers a simple, informal account of reversed-phi and four-stroke

motion. According to Fourier theory, any visual stimulus can be considered the sum of a set of sinusoidal grating components each having a specific spatial frequency, contrast, and spatial phase (position). When a photographic image is reversed in contrast, each frequency component in its Fourier spectrum reverses in contrast (dark bars become bright and vice versa). Contrast reversal of a sine wave is equivalent to a shift in its spatial phase of 180°, or half a cycle. So if the wave displaces, say, to the right by +90° and simultaneously reverses in contrast, the resulting displacement corresponds to a total rightward shift of +270° or equivalently a leftward shift of –90°. Assuming that the direction of perceived motion is governed by the direction of shortest displacement, a rightward displacement accompanied by contrast reversal should appear to be a leftward displacement; hence contrast reversal would lead to reversed-phi and four-stroke apparent motion. Anstis and Rogers (1986) did not specify the process that computes motion direction from phase shifts in Fourier components, arguing that almost any model that successfully detects real motion will sense motion in their stimuli.

The fact that stroke-based motion illusions elicit strong MAEs indicates that they stimulate the specialized neural motion sensors that are universally believed to underlie the MAE (see Mather, Pavan, Campana, & Casco, 2009). Adelson and Bergen (1985) proposed a detailed computational model of motion detection based on spatiotemporal energy (Fourier spatial frequency components that change phase over time). They showed that the output of their model offers a good account of reverse-phi effects (though Bours, Kroes, & Lankheet, 2007, presented a different theoretical perspective on reversed phi). The spatiotemporal energy model has since become the dominant model of early motion detection in the human visual system, supported by both psychophysical and physiological data (see Emerson, Bergen, & Adelson, 1992; Georgeson & Scott-Samuel, 1999).

Sensitivity to rapid temporal change is an essential characteristic of motion sensors, allowing them to respond to high velocities of movement across the retina. In order to achieve this sensitivity, the temporal response of motion sensors to a brief flash of light (known as the temporal impulse response [TIR]) is thought to contain two phases, a positive phase and a negative phase, as illustrated in Figure IV.73-3. In a two-frame display such as that shown in Figure IV.73-1a, with each frame typically lasting 30 to 40 msec with no IFI, the sensor's biphasic TIR drives down sensor response to the first frame in time for the appearance of the second frame, so the sensor can detect each frame separately and therefore encode the direction of motion.

Informally, the biphasic TIR in Figure IV.73-3 can explain two-stroke motion as follows. During the IFI the visual response to the first frame of the stimulus sequence enters its negative phase, and as a result the neural representation of the first frame reverses in contrast. Motion sensors then combine the contrast-reversed internal representation of the first frame with the initially positive-contrast representation of the second frame. The contrast-reversed internal representation shifts the spatial phase of all frequency components and thus reverses the signaled direction. Whereas in the case of four-stroke motion contrast reversal is physically applied to the stimulus before presentation, in two-stroke motion the contrast reversal occurs inside the visual system, as a consequence of the biphasic TIR of motion sensors.

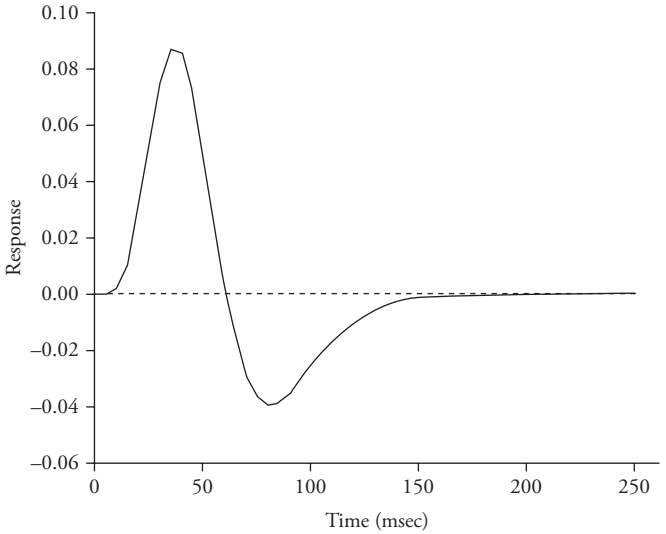


Figure IV.73-3. A biphasic TIR of the kind used in models of temporal vision and motion detection. The response to a brief flash of light at time zero reaches a positive peak within about 50 msec after the onset of the bright flash, and a negative peak within about 100 msec after the flash. (Figures by George Mather.)

To formally test this explanation of two-stroke motion, Mather and Challinor (2010) and Challinor and Mather (2010) implemented Adelson and Bergen's (1985) spatiotemporal energy model and applied it to the two-stroke stimulus. The model was found to provide a very good account of two-stroke motion, provided that the parameters of its TIR filters were selected appropriately. The dashed lines in Figure IV.73-2 show the output of the model as a function of IFI, superimposed on data from MAE experiments using the same IFIs. In bright conditions, the best-fitting TIR filters had a center temporal frequency of between 4 Hz and 6 Hz, and in dim conditions they had a center temporal frequency of 2 to 4 Hz.

It was noted earlier that IFI-reversal and two-stroke motion are both sensitive to the luminance of the IFI; both motion effects are strong using mean luminance IFIs but weak using bright or dark IFIs (Mather & Challinor, 2009; Shioiri & Cavanagh, 1990). When the spatiotemporal energy model is run with parameters set to the standard values used in the literature, it predicts no effect of IFI luminance. Changes in the parameters of the temporal filter make no difference. However, the model also includes parameters for spatial receptive fields, which take the standard form of a Gabor function (a sinusoid multiplied by a Gaussian) or, equivalently, a difference of Gaussians (DoG) function. Figure IV.73-4 (top left) shows a cross-section of this standard receptive field alongside its spatial frequency response (top right). Mather and Challinor (2010) reported that a slight change in the spatial receptive field profile, as shown in Figure IV.73-4 (bottom left and right), is sufficient to predict the obtained effect of IFI luminance on two-stroke motion perception. The relevant change in receptive field profile involves a reduction in the width of the Gabor's Gaussian space constant from 0.6x the period of the Gabor's sinusoid to 0.44x the period (equivalently, the introduction of a slight imbalance of 0.9 between the DoG's excitatory and inhibitory Gaussians). This change

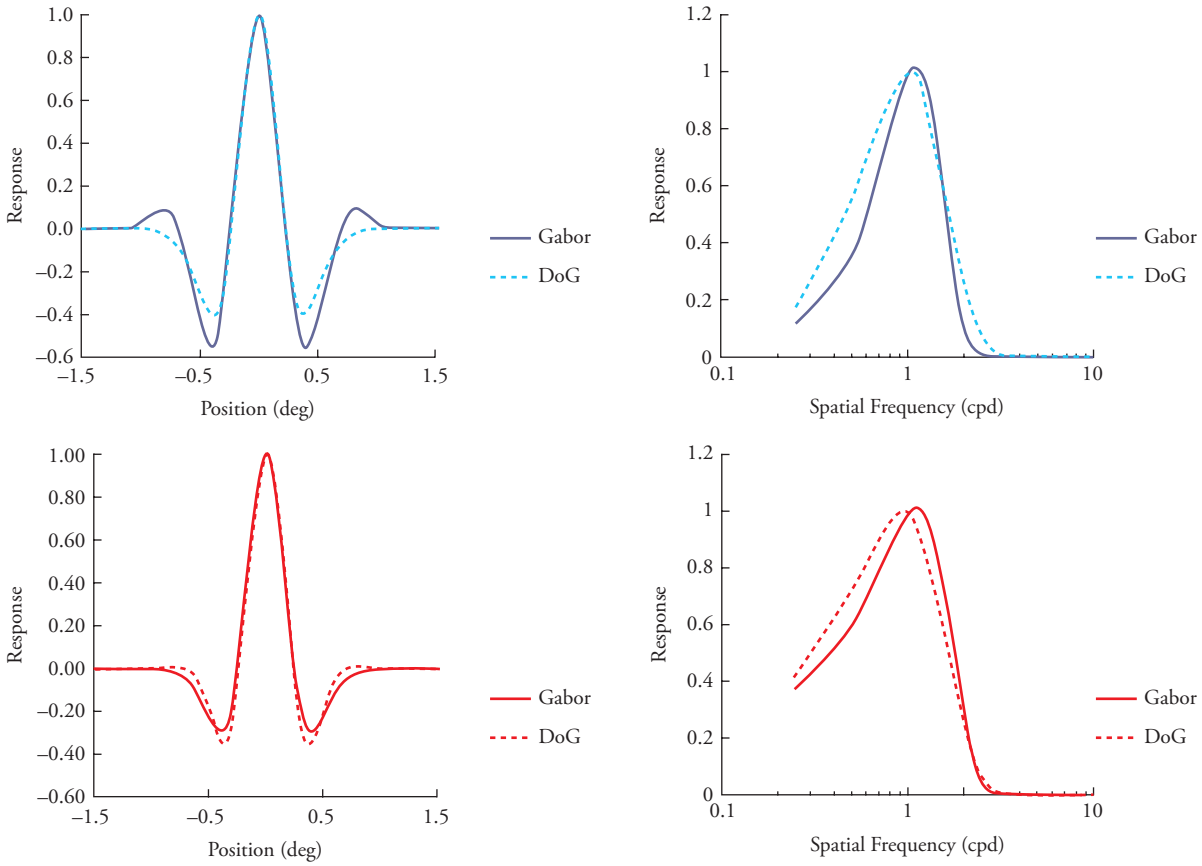


Figure IV.73-4. Spatial receptive fields used in the motion energy model tested against the two-stroke illusion. Top row: Standard Gabor and equivalent DoG filters used in published accounts of the model. The Gabor has a spatial frequency of 1.1 cpd and Gaussian space constant of 0.55°. The DoG function has an excitatory space constant of 0.25°, inhibitory space constant of 0.4°, and balance of 1.0 (relative area of excitatory and inhibitory Gaussians). Bottom row: Filter profiles required to predict the effect of IFI luminance on two-stroke motion. The bottom Gabor differs from the top Gabor in having a narrower Gaussian space constant of 0.4°, and the bottom DoG differs from the top DoG in having a balance factor of 0.9 (excitation slightly stronger than inhibition). (Figures by George Mather.)

introduces a slight DC bias into the spatial filter output, sufficient for the spatial receptive field to respond to changes in IFI luminance. Data from cat and monkey striate cortical receptive fields are consistent with this bias: Ringach (2002) reported that best-fitting Gabors had a Gaussian space constant of between 0.2x and 0.5x the Gabor's sinusoid.

In summary, two-stroke apparent motion is a robust, vivid impression of unidirectional motion created by alternating just two pattern frames, with a blank IFI at alternate frame transitions. It can be used to probe the properties of the underlying neural processes that encode retinal motion. A computational model of motion detection based on spatiotemporal energy can explain the effect, provided that its parameters are set to give a biphasic temporal response and a slightly imbalanced spatial response.

AUTHOR'S NOTE

For demonstrations of stroke-based motion illusions, see http://www.georgemather.com/MotionDemos/TwostrokeMP4.html. For a detailed description and Matlab code for Adelson and Bergen's (1985) spatiotemporal energy model of motion detection, see http://www.georgemather.com/Model.html.

REFERENCES

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, *2*(2), 284–299.

Anstis, S. M. (1970). Phi movement as a subtraction process. *Vision Research*, *10*, 1411–1430.

Anstis, S. M., & Rogers, B. J. (1975). Illusory reversal of visual depth and movement during changes of contrast. *Vision Research*, *15*, 957–961.

Anstis, S. M., & Rogers, B. J. (1986). Illusory continuous motion from oscillating positive–negative patterns: Implications for motion perception. *Perception*, *15*, 627–640.

Bours, R. J. E., Kroes, M., & Lankheet, M. (2007). The parallel between reverse-phi and motion aftereffects. *Journal of Vision*, *7*(8), 1–10.

Challinor, K., & Mather, G. (2010). A motion-energy model predicts the direction discrimination and MAE duration of two-stroke apparent motion at high and low retinal illuminance. *Vision Research*, *50*, 1109–1116.

Emerson, R. C., Bergen, J. R., & Adelson, E. H. (1992). Directionally selective complex cells and the computation of motion energy in cat visual cortex. *Vision Research*, *32*(2), 203–218.

Gegenfurtner, K. R., Mayser, H. M., & Sharpe, L. T. (2000). Motion perception at scotopic light levels. *Journal of the Optical Society of America A*, *17*(9), 1505–1515.

Georgeson, M. A., & Scott-Samuel, N. E. (1999). Motion contrast: A new metric for direction discrimination. *Vision Research*, *39*, 4393–4402.

Kolers, P. A. (1972) *Aspects of motion perception*. Oxford: Pergamon.

Mather, G. (2006). Two-stroke: A new illusion of visual motion based on the time course of neural responses in the human visual system. *Vision Research*, *46*, 2015–2018.

Mather, G., & Challinor, K. (2010) Psychophysical tests of the motion energy model. *Perception*, *39*, S152.

Mather, G., Pavan, A., Campana, G., & Casco, C. (2009). The motion after-effect reloaded. *Trends in Cognitive Sciences*, *12*(12), 481–487.

Ringach, D. L. (2002). Spatial structure and symmetry of simple-cell receptive fields in macaque primary visual cortex. *Journal of Neurophysiology*, *88*(1), 455–463.

Shioiri, S., & Cavanagh, P. (1990). ISI produces reverse apparent motion. *Vision Research*, *30*, 757–768.

Strout, J. J., Pantle, A., & Mills, S. L. (1994). An energy model of interframe interval effects in single-step apparent motion. *Vision Research*, *34*, 3223–3240.

Takeuchi, T., & De Valois, K. K. (1997). Motion-reversal reveals two motion mechanisms functioning in scotopic vision. *Vision Research*, *37*, 745–755.

Chapter 74
On the Pinna Illusion

Baingio Pinna

INTRODUCTION

The Pinna illusion is the first case of visual illusion showing a rotating motion effect. Figure IV.74-1a shows a pattern of grey square-shaped checks, made up of two white and two black edges each, on a grey background. The squares, all with the same width, length, and orientation in relation to the center of their circular arrangements, are grouped in two concentric rings due to the Gestalt proximity principle. The difference between the checks of the two rings is in the relative position of their black and white edges along the vertexes: whereas the edge polarity of the checks in the surrounding ring is from black to white (going from left to right), in the inner ring it is from white to black. More precisely, the checks of the two rings show reversal of the vertex orientation and, consequently, opposite inclination of the virtual/implicit diagonal orientation polarity, obtained by joining the two vertexes where black and white lines meet (Pinna, 1990; Pinna & Brelstaff, 2000).

By slowly moving the head toward the figure with the gaze fixed in the central black dot, the inner ring of squares starts to rotate counterclockwise, while the outer ring rotates clockwise. The direction of rotation reverses when the head moves away from the stimulus pattern; that is, the inner ring rotates now clockwise and the outer one counterclockwise. The apparent motion is instantaneous and perpendicular to the direction of the true motion, and the speed of the apparent rotation is proportional to the one of the motion imparted by the observer.

A dynamic simulation of the action of moving toward and away from the figure by physically expanding and contracting Figure IV.74-1a is represented by Video IV.74-1. When the observer keeps the gaze fixed on the center and the head stationary, the two concentric rings of squares are perceived to counterrotate as previously described. By physically rotating clockwise the stimulus pattern of Figure IV.74-1a, as shown dynamically in Video IV.74-2, the inner ring is seen to contract while the outer one is perceived to expand. The opposite result is perceived when the stimulus pattern is rotated counterclockwise (not illustrated).

ILLUSION AND ILLUSORINESS IN THE PINNA ILLUSION

These results suggest that the Pinna illusion subsumes at least two main phenomena: the illusion of rotation and the illusoriness (Pinna, 1992, 2012) of the effect. In Figure IV.74-1a, the rotation is perceived "true" and "real" (dependent on the observer motion) and, at the same time, "illusory." The main source of illusoriness is the fact that the true motion of the observer is perceived as occurring simultaneously or as "causing" the counterrotation of the circles of squares. More particularly, the observer perceives him or herself as the cause of the motion of something that is seen at the same time as still. The paradoxical impression of illusoriness is related to the fact that the motion of the observer is not usually followed by the motion of the objects of the real world as it is instead the case of Figure IV.74-1a. Therefore, the self-motion and the proprioception are not discounted but appear related to the perceived counterrotation of the squares. As a consequence, observers usually move their heads back and forth many times to clarify the strong connection between the two motions. The illusoriness is consequently demonstrated by the astonishment based on the question "Why do the rings move?" and is determined by the perceived relation of cause and effect between the observer and the stimulus pattern.

MAIN PROPERTIES OF THE ILLUSION

Seeing Peripherally

The apparent rotation is perceived in peripheral vision. In Figure IV.74-1a, looming the head in and out when the gaze is fixed not on the center but on any square, the counterrotation effect is destroyed and thus the fixed square appears motionless. This is due to the motion detectors in the fovea that are less sensitive to this type of stimuli than in the periphery. Moreover, during peripheral viewing the spatial square shape of each check should be

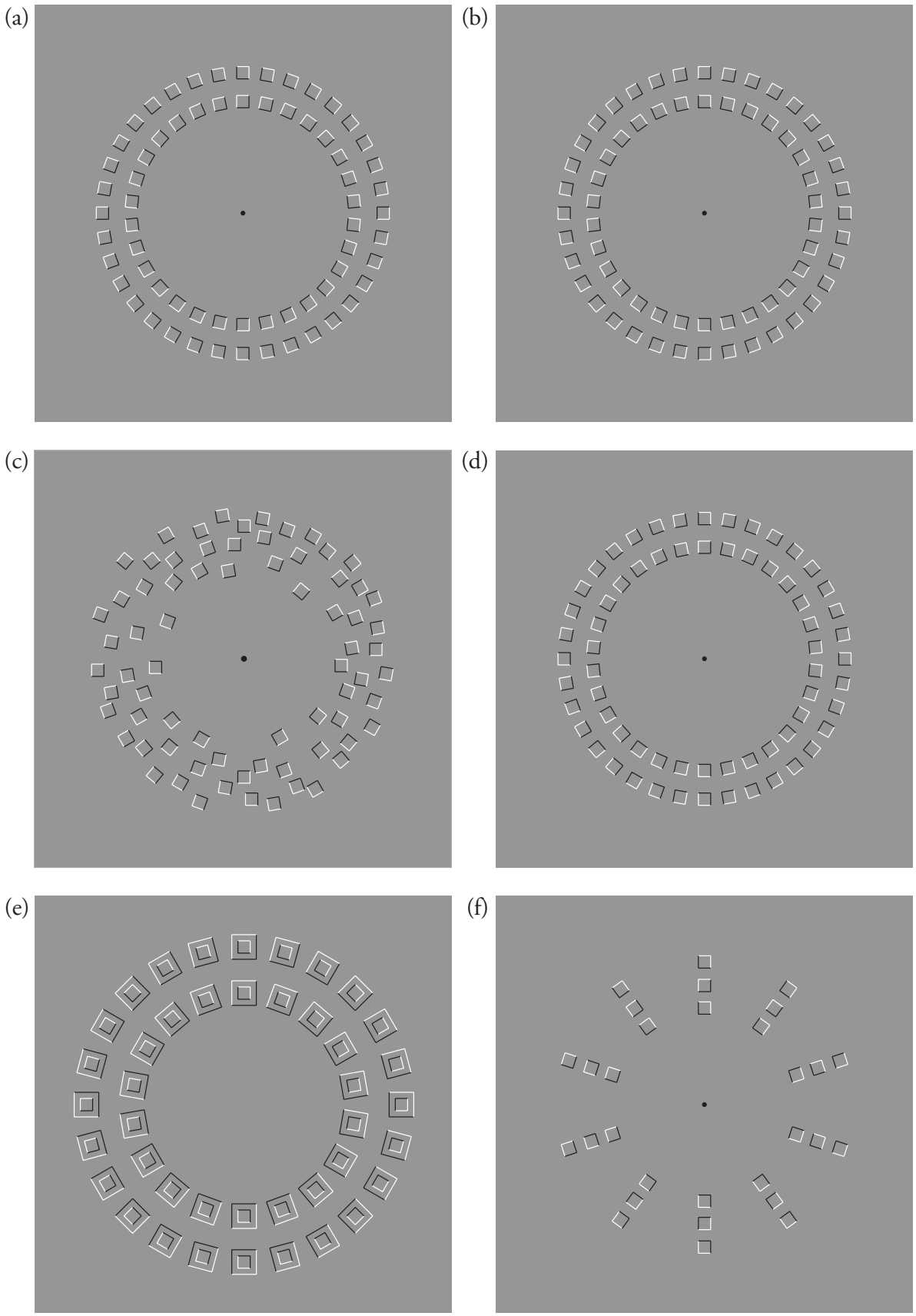
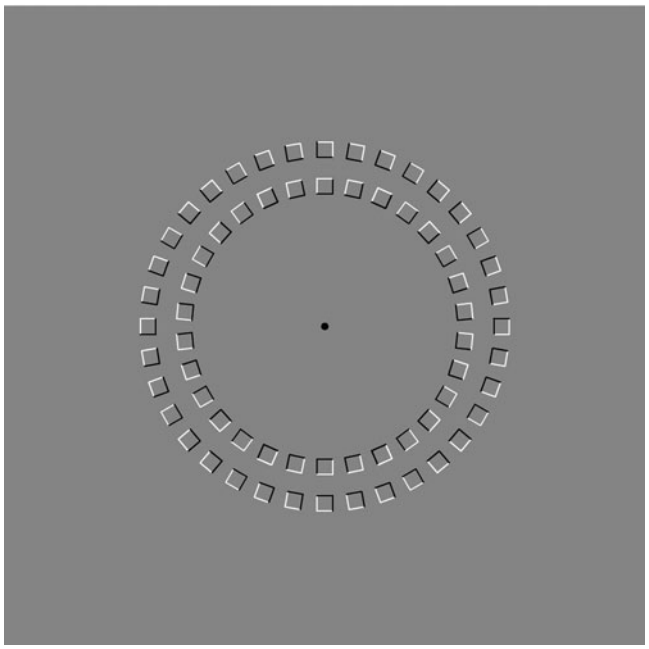
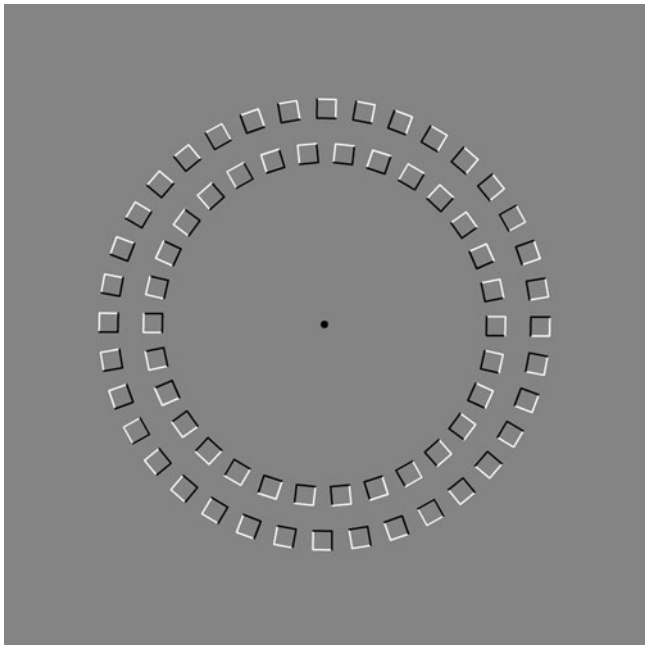


Figure IV.74-1. Examples of the Pinna illusion. (a) By slowly moving the head towards the figure with the gaze fixed in the central black dot, the inner ring of squares starts to rotate counter-clockwise, while the outer ring rotates clockwise. (b) By reversing the implicit diagonal orientation polarity of each check, the perceived rotation reverses accordingly. (c) The alignment of the checks is not necessary to elicit the illusory rotation. (d, e) The illusory counter-rotation is nulled or (f) reduced. (Figure by Dr. Baingio Pinna's creations.)



Video IV.74-1. Video by Dr. Baingio Pinna's creations.



Video IV.74-2. Video by Dr. Baingio Pinna's creations.

blurred; therefore the main motion cues derive from the whole check element and not from the constituent and elementary line segments.

Reversing the Implicit Diagonals

As illustrated in Figure IV.74-1b, by reversing the implicit diagonal orientation polarity of each check, rendered by joining the two vertexes where the black and white segments meet, the perceived rotation reverses accordingly. This result demonstrates that it is the difference in slant of the implicit diagonals of the squares that provides the directional cues for biasing the global motion vector thus affecting the perceived motion direction.

Misaligning Implicit Diagonals

The alignment of the checks is not necessary to elicit the illusory rotation. As shown in Figure IV.74-1c, when the checks are randomly shuffled and misaligned in concentric circles, the apparent rotation arises and persists.

This result suggests that the organization of the squares does not require oriented Fourier components that may serve as a secondary feature to the motion system. Therefore, the apparent motion is individually assigned to each square on the basis of the arrangement of the black and white sides; afterward the perception of the counterrotation depends on the grouping and segregation of the two populations of squares according to the Gestalt principle of common fate.

Alternating Implicit Diagonals

In Figure IV.74-1a, the two kinds of implicit orientation polarities are grouped circularly, each in a different radius (i.e., in different locations of the stimulus, one inside and the other outside). In terms of Gestalt grouping it means that the common fate is synergistic with the good continuation and the relative position (inside vs. outside), thus creating the smoothest continuation. The question is now "What happens when the orientation polarity of the checks within each ring is alternated?" therefore breaking the whole direction and smoothest continuation of the individual assignment of motion. As demonstrated in Figure IV.74-1d, the illusory counterrotation is nulled. A similar result is perceived in Figure IV.74-1e, where the orientation polarities of two checks, one included within the other and arranged in two rings, are opposite.

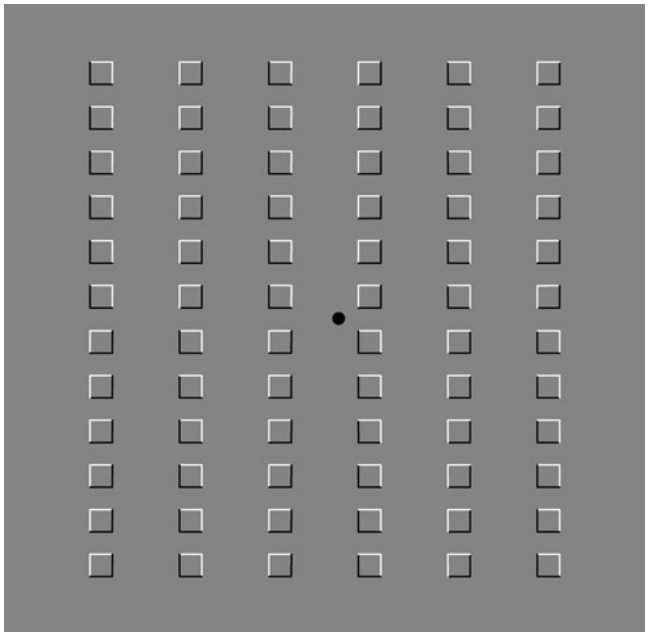
Just as Figure IV.74-1a demonstrates the grouping of motion signals that belong to the same global direction, the results of Figures IV.74-1d and IV.74-1e demonstrate the mutual inhibition of opposite-orientation polarities, thus corroborating the hypothesis of short- and long-range interactions among motion vectors.

In Figure IV.74-1f, the grouping of the squares by proximity is not circularly (as in Fig. IV.74-1a) but radially oriented, while the single squares maintain their alternated orientation polarities from the inside to the outside of the stimulus patterns similarly to Figure IV.74-1a (Pinna & Gregory, 2009). Under these conditions, the counterrotating effect is strongly reduced, and a radial waving or twisting along the grouping is perceived.

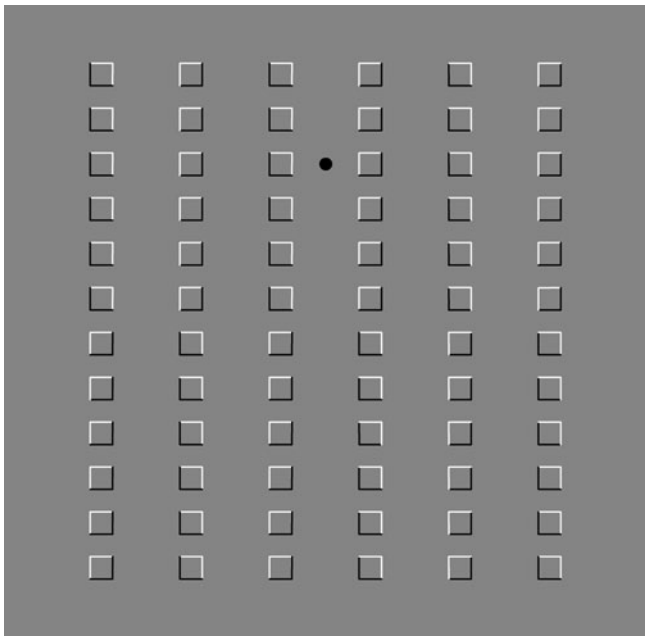
Grouping Implicit Diagonals in Rows and Columns

In the previous figures, the apparent motion is elicited by an optical flow field corresponding to radial expansion or contraction (Gibson, 1979). In Videos IV.74-3, IV.74-4, and IV.74-5, the apparent motion arises by a parallel vector field in a pattern of checks arranged in rows and columns (cf. Pinna, 1990).

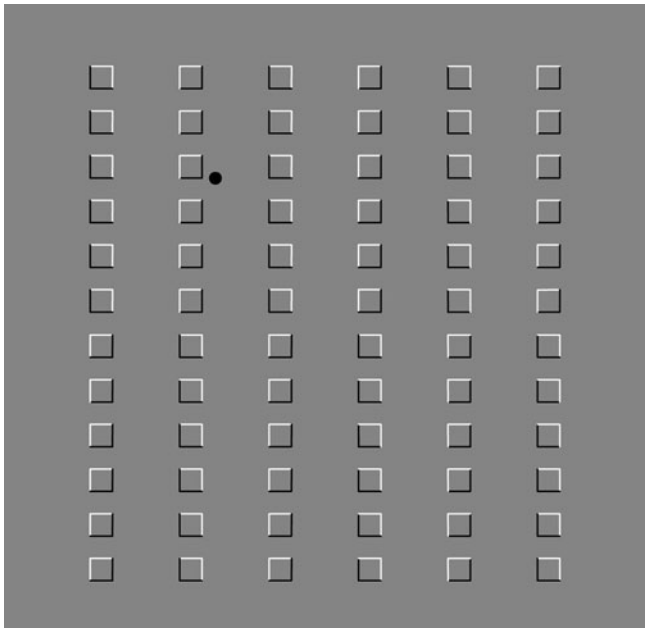
In Video IV.74-3, the alternation of black and white edges of the checks in between columns and from the bottom half to the top half of the pattern of checks produces a strong



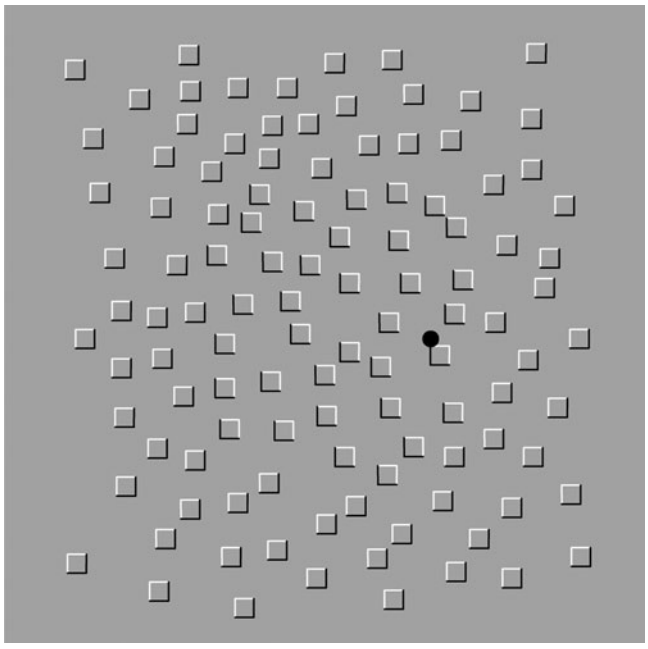
Video IV.74-3. Video by Dr. Baingio Pinna's creations.



Video IV.74-4. Video by Dr. Baingio Pinna's creations.



Video IV.74-5. Video by Dr. Baingio Pinna's creations.



Video IV.74-6. Video by Dr. Baingio Pinna's creations.

vertical sliding motion when the gaze is fixed on the horizontally moving dot: the columns are perceived as moving up and down, alternately in the two top and bottom halves. Apparent moving convergence and divergence of the columns is also perceived.

In Video IV.74-4, as the gaze follows the vertically moving dot, the columns of checks now appear to move left and right, alternately in the two halves. Meanwhile, in Video IV.74-5, the apparent motion is strongly reduced or annulled by following with the gaze the diagonally moving dot.

In Video IV.74-6, by keeping the gaze on the moving dot, the central irregular cluster of checks immediately segregates from the surrounding ones. This is a further demonstration of the role of the Gestalt factor of common fate and that the opposite orientation of the implicit diagonals provides the directional cues for biasing the global motion vector (Gurnsey & Pagé; 2006; Gurnsey, Sally, Potechin, & Mancini, 2002; Morgan, 2002; Pinna & Brelstaff, 2000; Pinna & Spillmann, 2002, 2005).

Implicit Versus Explicit Diagonals

Within the kinds of check elements used here, there are two different orientation cues that can instill polarity to

each basic element: (a) the orientation of the invisible diagonal joining the two vertexes where the black and white sides meet—implicit orientation (demonstrated in the previous figures), and (b) the orientation of the checks themselves—explicit orientation.

Figures IV.74-2a and IV.74-2b demonstrate that the explicit orientation (tilt) of the squares can be either synergistically summed (Fig. IV.74-2a) or pitted against (subtracted from) the implicit orientation (Fig. IV.74-2b). By zooming the figures while keeping the gaze fixed in its

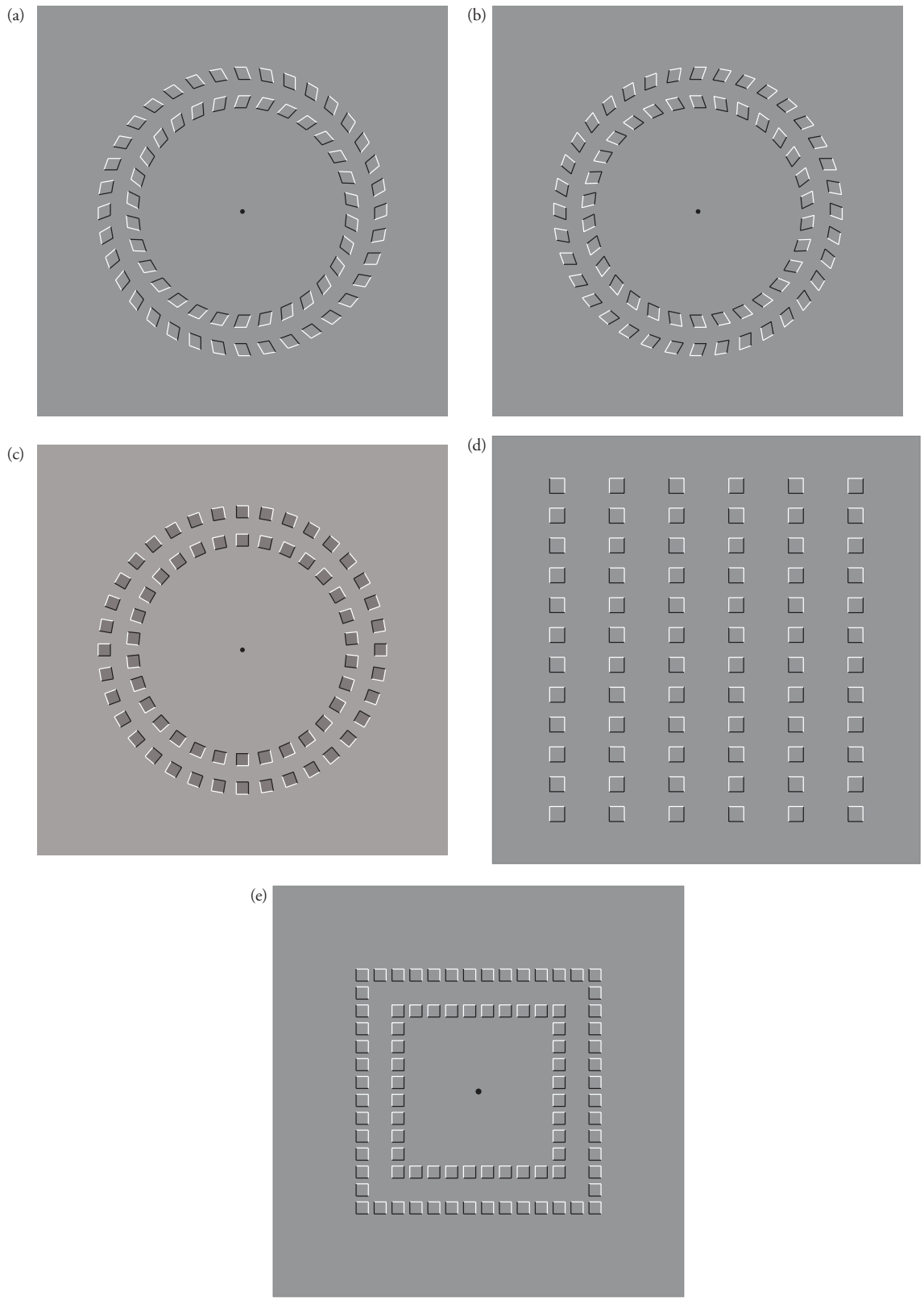


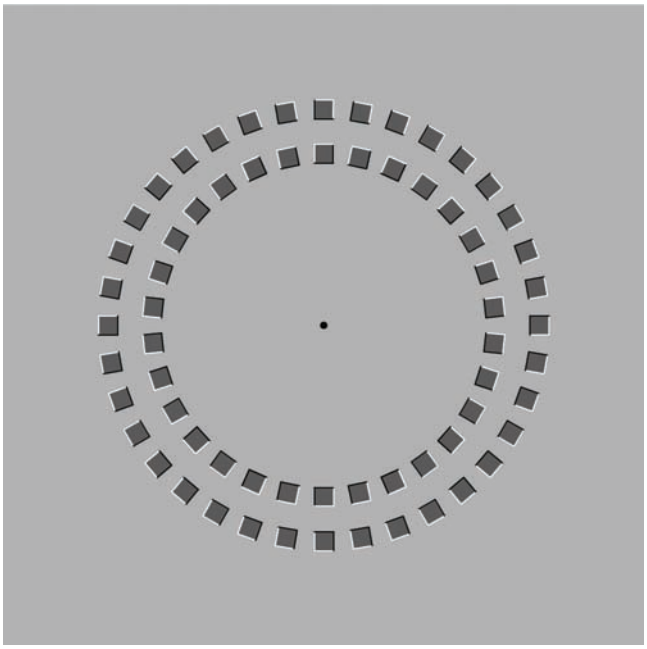
Figure IV.74-2. (a, b, c) The tilt of the squares, the asymmetric luminance profile can respectively increase or decrease the effect of counter-rotation. (d, e) Alternated implicit diagonals, viewed peripherally, induce tilt distortions of each check (d) and of the whole square (e) that appear deformed as rhombic shapes. (Figure by Dr. Baingio Pinna's creations.)

center, the effect of counterrotation can be either increased (Fig. IV.74-2a), if compared with one of Figure IV.74-1, or decreased (Fig. IV.74-2b).

Asymmetric Luminance Profile

A further directional cue for biasing motion vectors inherent in the previous figures is the asymmetric luminance profile. In Figure IV.74-2c, by keeping the gaze fixed in the center while the head is moved toward the figure, the inner ring of squares appears to rotate counterclockwise and the outer one clockwise as demonstrated in Figure IV.74-1a. However, when the head is moved away from the stimulus pattern, the squares of the inner and outer rings do not appear to rotate (or just barely), differently from what is expected from Figure IV.74-1a. This clear asymmetry in the strength of the apparent rotation depends on the asymmetric luminance profile of the black bars, the dark grey within each square, the white bars, and the light grey background. This luminance profile represents a strong directional bias and a different kind of explicit polarity in the direction of the minimum luminance intensity change (i.e., toward the white bars). Other effects are based on this cue, such as the staircase illusion (Fraser & Wilcox, 1979) and the peripheral drift illusion (Faubert & Herbert, 1999; Kitaoka & Ashida, 2003).

Just as the explicit orientations of the checks of Figure IV.74-2b can be synergistically added to or subtracted from the implicit orientations, the effect obtained in Figure IV.74-2c is related to the summation (when the head is moved toward the figure) and to the subtraction (when the head is moved away) of the implicit orientations from the asymmetric luminance profile. In Video IV.74-7, the summation/subtraction effects, alternated in time in two photograms, produce counterclockwise rotation and expansion alternated in a loop with clockwise rotation and contraction.



Video IV.74-7. Video by Dr. Baingio Pinna's creations.

PHENOMENA RELATED TO THE PINNA ILLUSION

The Rhombic Squares

In Figure IV.74-2d, alternated implicit diagonals grouped in columns, when viewed peripherally, induce tilt distortions of each check that appears deformed as a rhombic shape in the same direction as the implicit diagonal. The rhombic distortion is perceived even more strongly by slowly moving the gaze along the stimulus pattern. In Figure IV.74-2e, the rhombic distortion can also involve the whole square made up of small squares; that is, the implicit diagonals of the small checks produce a holistic distortion of the global square in the same direction as the implicit diagonals.

Illusory Convergence/Divergence

In Figure IV.74-3a, the alternation of the implicit diagonals of squares with black and white sides in the two vertical halves of the pattern, grouped in columns, deforms the geometrical vertical parallelism of the squares that appear converging and diverging. As demonstrated by comparing Figures IV.74-3a and IV.74-3b, the converging–diverging effect is the opposite of the Zöllner illusion (Zöllner, 1860). This result is unexpected if the implicit diagonals of the checks are considered equivalent to the intersecting parallel segments of the Zöllner illusion.

If this is true, then the inverted Zöllner illusion due to implicit diagonals can be enhanced or annulled by adding explicit diagonals, as shown in Figures IV.74-3c and IV.74-3d. It is worthwhile noticing that, while the converging–diverging phenomenon is enhanced, the illusory sliding motion is annulled (Fig. IV.74-3c), and, vice versa, when the converging–diverging phenomenon is annulled, the illusory sliding motion is enhanced (Fig. IV.74-3d).

In Figure IV.74-3e, the strength of the converging–diverging effect of Figure IV.74-3a is enhanced by reversing the black and white sides in alternated squares along the columns. This is due to the summation of the direction of the implicit diagonals with the direction of the black and white arrangement of sides that is synergistic with the one of the implicit diagonals.

Phenomenally, all these effects are better perceived in peripheral or at a global viewing, similarly to the Pinna illusion, while the converging–diverging phenomenon of the Zöllner illusion is also perceived foveally and locally.

Intertwining and Spiral Effects

By geometrically arranging the squares in two concentric circles as shown in Figure IV.74-4a, where the implicit diagonals are alternated between two contiguous circles, the converging–diverging effect is now perceived as an intertwining phenomenon. When the implicit diagonals have the same orientation in between the two circles, the two rings tend to be placed in a spiral shape (Fig. IV.74-4b). Under these conditions, by zooming the figure while the gaze is kept fixed in its center, the two rings are perceived as rotating in the same direction. The effects illustrated in Figures IV.74-4a and IV.74-4b are similar to the intertwining effect illustrated in Figures IV.74-4c and IV.74-4d

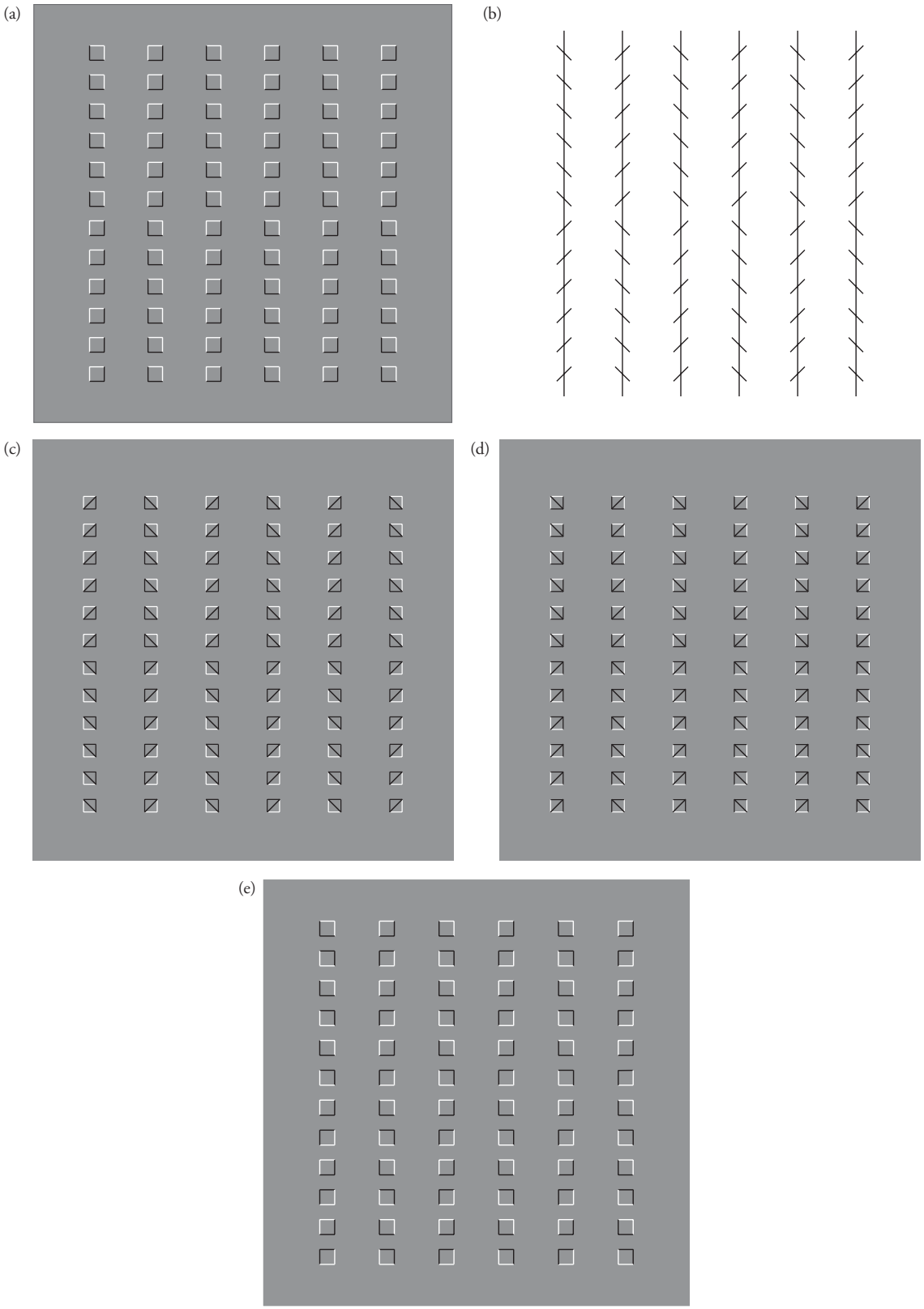


Figure IV.74-3. (a) The alternation of the implicit diagonals deforms the geometrical vertical parallelism of the squares that appear converging and diverging. The converging-diverging effect is opposite to the Zöllner illusion (b-e). (Figure by Dr. Baingio Pinna's creations.)

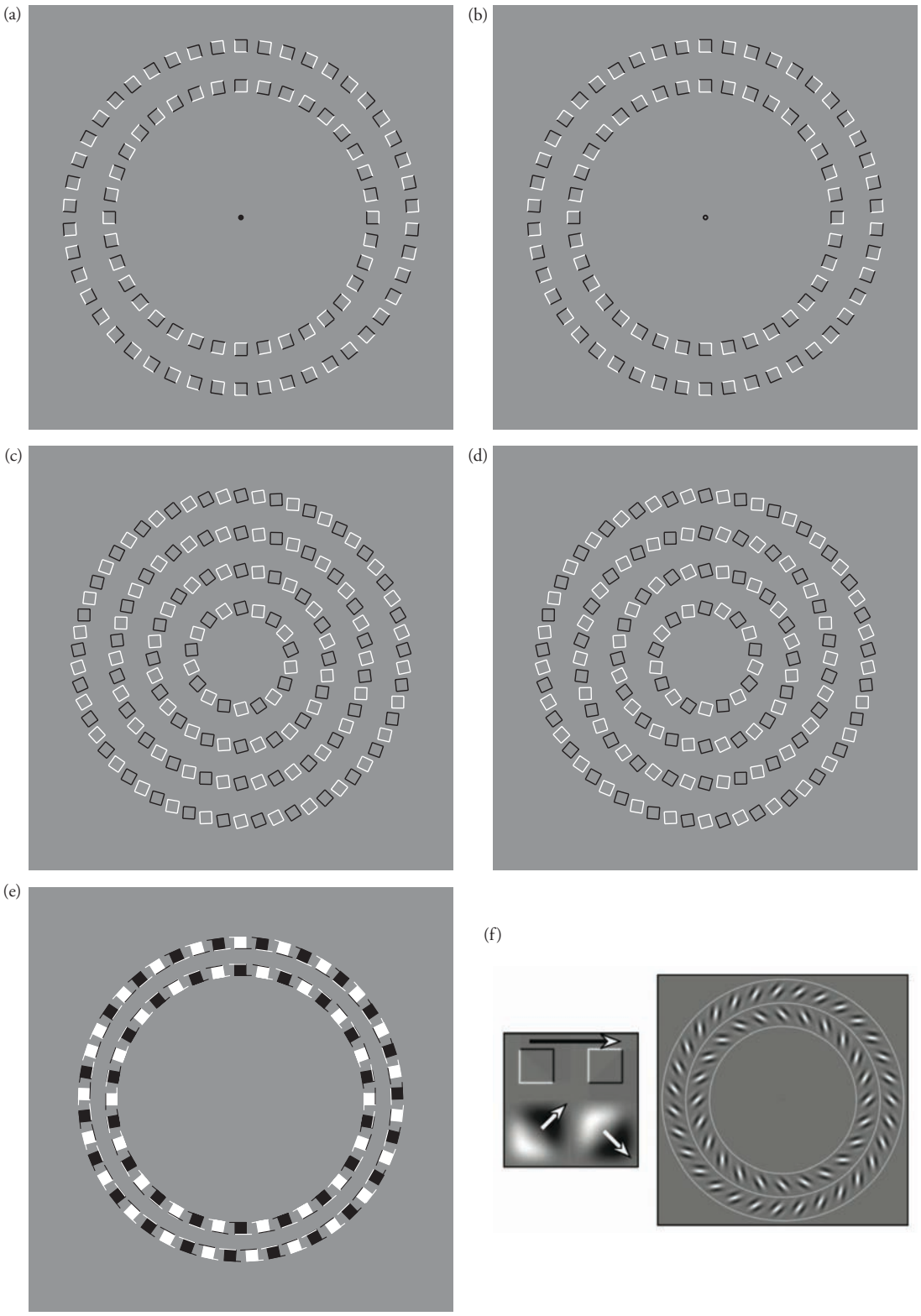


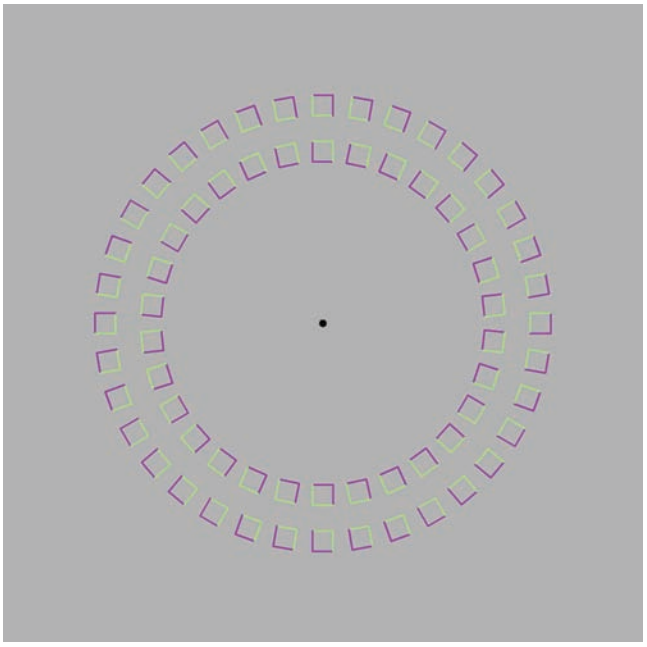
Figure IV.74-4. (a,b) Intertwining and spiral effects due to the implicit diagonals. (c-e) When the implicit diagonals are replaced by tilted squares, similar intertwining and spiral effects emerge but in the opposite direction. (f) By removing the high frequencies from the checks, the strength of the illusion persists or is even enhanced. (Figure by Dr. Baingio Pinna's creations.)

and to the spiral effect shown in Figure IV.74-4e (Pinna & Gregory, 2002).

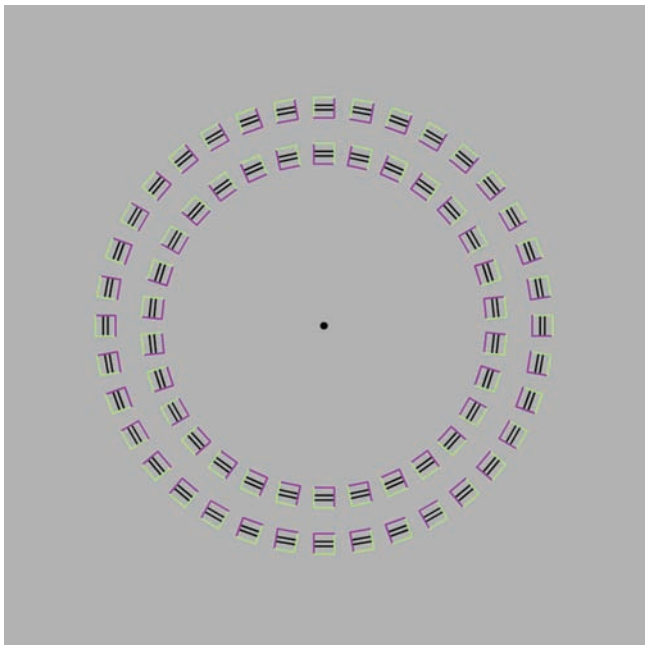
The Role of Explicit Directions During Luminance Modulation

The role of the explicit directions, shown in Figures IV.74-3a, IV.74-3c, and IV.74-3d, is clearly demonstrated in another related phenomenon perceived under the luminance contrast modulation. In the two continuously alternated frames of Video IV.74-8, the green edges get brighter and darker, while the purple ones become darker and brighter. As expected, by fixing the center of the stimulus, the luminance modulation is perceived as a contrast pulsation. A very weak motion effect can also be perceived. However, by adding two parallel segments to each square, tangent to the virtual circumferences of the annuli (see Video IV.74-9), the two annuli appear to be ticking counterclockwise and clockwise similarly to a clock. Some weak expansion and contraction can also be perceived. By placing the segments perpendicularly, an alternated expansion and contraction is perceived (see Video IV.74-10); that is, while one annulus expands the other contracts.

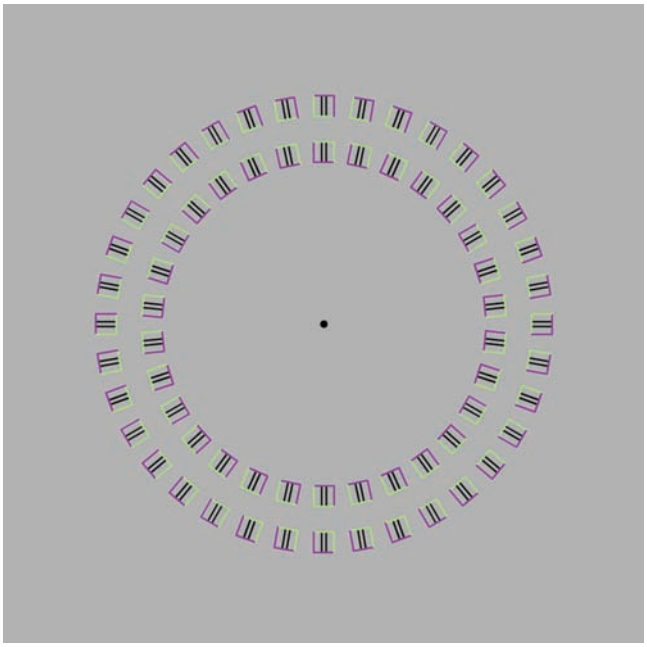
These results demonstrate that the explicit orientations polarize the motion vectors due to the contrast modulation. The synergistic organization of explicit orientations and contrast modulation is further demonstrated in Video IV.74-11, where a clear ticking is perceived. In Video IV.74-12, modulating dots, invisible in peripheral viewing (i.e., when the gaze is fixed on the center of the figure) instill a ticking motion into the black edges. Other results are demonstrated in Video IV.74-13 and Video IV.74-14. In Video IV.74-15, the illusory rotation is not perceived as a back and forth ticking, as expected on the basis of the contrast modulation of the juxtaposed edges, but as a continuous counterclockwise ticking. These results are related to the "phenomenal phenomena" described by Gregory and Heard (1972), the reverse phi and four-stroke motion (Anstis & Rogers, 1975, 1986), and the visual illusions based on single-field contrast asynchronies (Shapiro, Charles, & Shear-Heyman, 2005).



Video IV.74-8. Video by Dr. Baingio Pinna's creations.



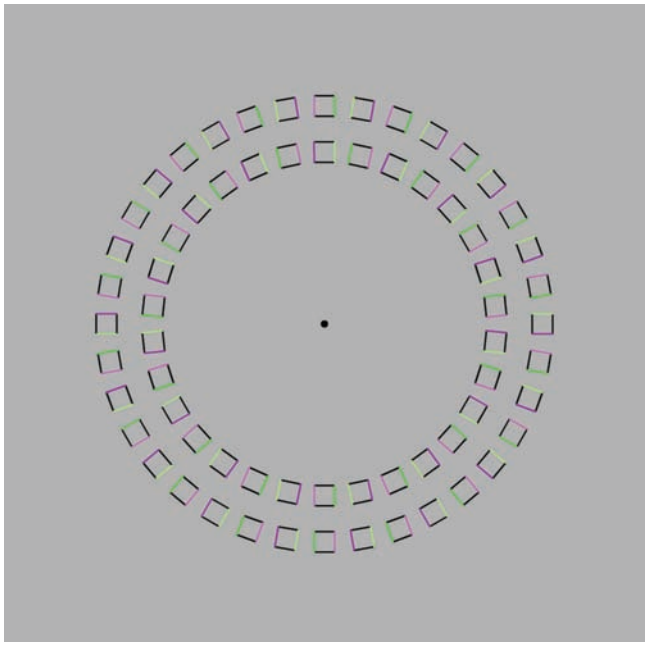
Video IV.74-9. Video by Dr. Baingio Pinna's creations.



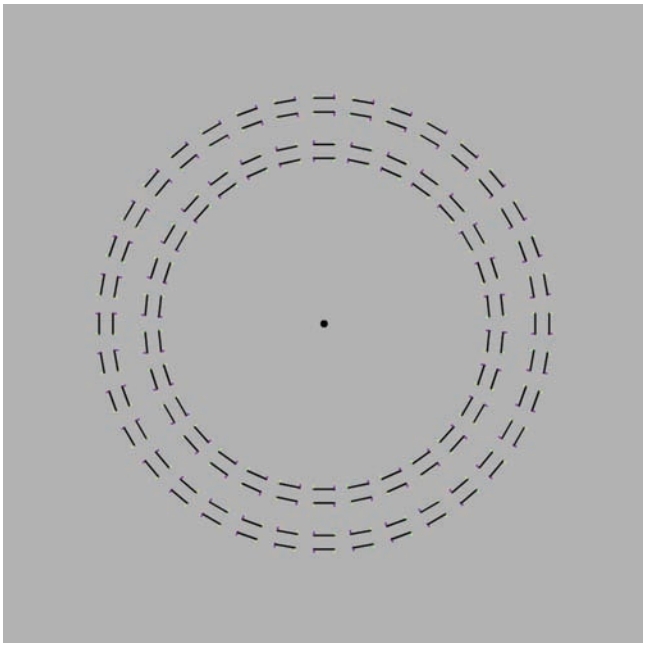
Video IV.74-10. Video by Dr. Baingio Pinna's creations.

NEURAL MECHANISMS UNDERLYING THE PINNA ILLUSION

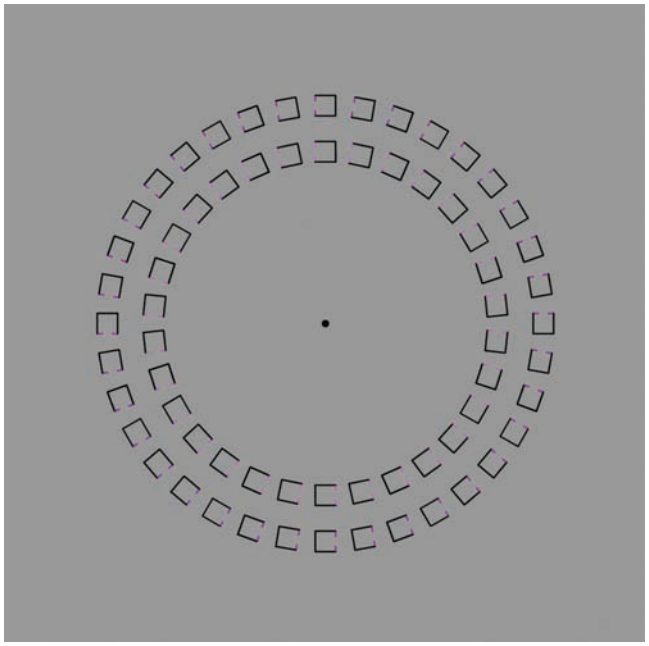
The interpretation of the motion illusion depends on a step where image features such as edges, intersections of edges, black and white edges, and local image movement are derived. These features contain many sources of noise



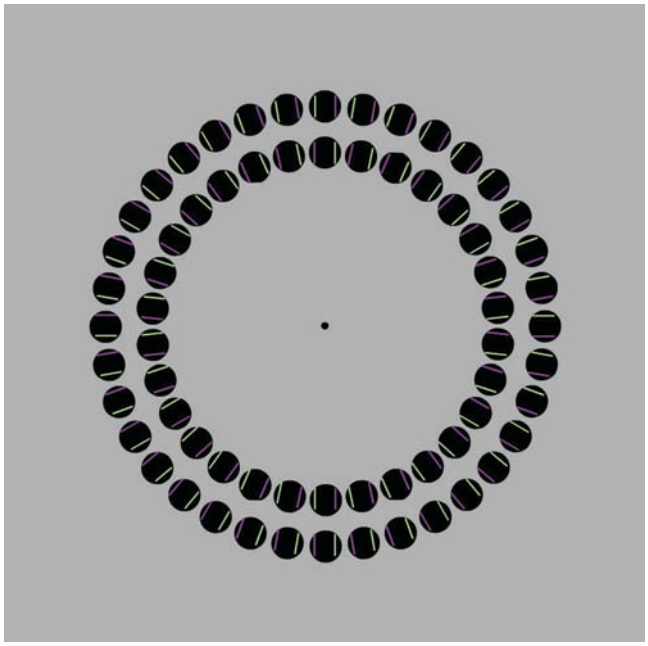
Video IV.74-11. Video by Dr. Baingio Pinna's creations.



Video IV.74-13. Video by Dr. Baingio Pinna's creations.



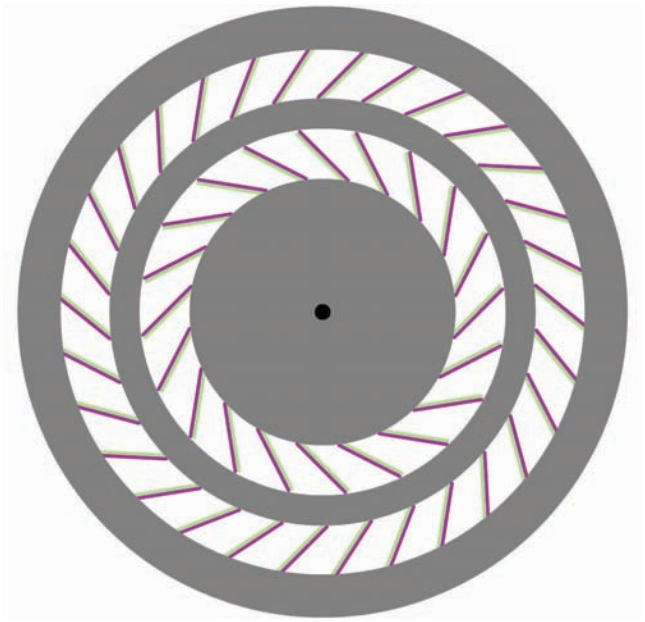
Video IV.74-12. Video by Dr. Baingio Pinna's creations.



Video IV.74-14. Video by Dr. Baingio Pinna's creations.

or uncertainty that can cause bias. As a result, the locations of features can be perceived erroneously and the appearance of the patterns can be altered (see Fermüller & Malm, 2004; Fermüller, Pless, & Aloimonos, 2000; Weiss & Fleet, 2002; Weiss, Simoncelli, & Adelson, 2002). The basic attributes underlying the Pinna illusion are the implicit orientation polarities (diagonals) in the check micropatterns, namely the low-frequency luminance gradients, not the high-frequency black and white edges of the squares. The implicit diagonals suggest that the illusion can be accounted for in terms of orthogonal biases (Grossberg, Mingolla, & Viswanathan, 2001; Gurnsey & Pagé, 2006; Gurnsey et al., 2002; Mather, 2000; Pinna & Brelstaff, 2000; Pinna & Spillmann, 2005).

In Figure IV.74-4f (left), the two bottom micropatterns are the blurred version of the two above, obtained by removing the high frequencies from the checks. Translating the micropatterns to the right causes them to stimulate most strongly neurons selective for directions as indicated by the white arrows. Thus the estimated flow vectors of Figure IV.74-1a are biased in the clockwise and in the counterclockwise directions. The role of low-frequency luminance gradients is demonstrated by replacing the micropatterns with Gabor patches (Bayerl & Neumann, 2002; Gurnsey



Video IV.74-15. Video by Dr. Baingio Pinna's creations.

& Pagé, 2006; Gurnsey et al., 2002; Morgan, 2002). In this case, the strength of the illusion persists or is even enhanced (see Fig. IV.74-4f, right).

The illusion can be explained by the response of direction-selective neurons at the earliest cortical stage (area V1) of visual processing. These neurons signal the speed with which a line with its preferred orientation moves through its receptive field similarly to what happens in the aperture effect (cf. Nakayama & Silverman, 1988), where a moving straight line seen through an aperture is perceived to move only along the direction of its normal pathway. While this explains how each individual check receives a local illusory motion signal, the whole illusory rotational motion can be thought to be sensed by the higher cortical area such as MT (medium-scale motion analysis, inhibition of opponent directions) and dorsal MST (large-scale motion analysis, directional decomposition), which collates all the signals provided by the local motion checks.

The Pinna illusion represents an opportunity within the context of vision science and cognitive neuroscience (Gazzaniga, 2004; Purves & Lotto, 2003) because it reveals two issues that challenge theories and deserve to be further investigated. The first issue is related to the role of the observer's motion. In everyday life the pattern of motion (optical flow; Gibson, 1979) of objects, surfaces, and edges in a visual scene is not perceived as an illusory motion of the objects but only as the motion of the observer. In other words, while in a natural scene the final motion outcome of the objects is compensated by messages from other sensory systems (e.g., proprioception), in the Pinna illusion it is the observer's head movement that is perceived as the cause that instills the perceived motion. The second issue is related to the illusion of motion and the perceived illusoriness (Pinna, 1992, 2008, in press) of the apparent motion of a pattern seen, at the same time, as static.

The study of the relationship between illusion and illusoriness is a further challenge for neuroscientists (Pinna, in press). Both issues raised by the Pinna illusion can shed new light in the no-man's-land between "sensory" and "cognitive" processes not fully explored yet.

ACKNOWLEDGMENTS

This work was supported by Banco Fondazione di Sardegna - Rep. n. 2539. Prot. n. 23907 del 05/10/2015.

REFERENCES

Anstis, S. M., & Rogers, B. J. (1975). Illusory reversal of visual depth and movement during changes of contrast. *Vision Research*, *15*(8–9), 957–961.

Anstis, S. M., & Rogers, B. J. (1986). Illusory continuous motion from oscillating positive-negative patterns: Implications for motion perception. *Perception*, *15*(5), 627–640.

Bayerl, P., & Neumann, H. (2002). Cortical mechanisms of processing visual flow—Insights from the Pinna-Brelstaff illusion. In Dynamic Perception, RP Würtz, M. Lappe (eds.), In *Proceedings of the 4th Workshop: Dynamische Perzeption*, Bochum, AKA, 23–28.

Faubert, J., & Herbert, A. M. (1999). The peripheral drift illusion: A motion illusion in the visual periphery. *Perception*, *28*, 617–621.

Fermüller, C., & Malm, H. (2004). Uncertainty in visual processes predicts geometrical optical illusions. *Vision Research*, *44*, 727–749.

Fermüller, C., Pless, R., & Aloimonos, Y. (2000). The Ouchi illusion as an artifact of biased flow estimation. *Vision Research*, *40*, 77–96.

Fraser, A., & Wilcox, K. J. (1979). Perception of illusory movement. *Nature*, *281*, 565–566.

Gazzaniga, M. S. (2004). *The cognitive neurosciences III*. Cambridge, MA: MIT Press.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Gregory, R. L., & Heard P. (1972). Visual dissociations of movement, position, and stereo depth: Some phenomenal phenomena. *Quarterly Journal of Experimental Psychology*, *35A*, 217–237.

Grossberg, S., Mingolla, E., & Viswanathan, L. (2001). Neural dynamics of motion integration and segmentation within and across apertures. *Vision Research*, *41*, 2521–2553.

Gurnsey, R., & Pagé, G. (2006). Effects of local and global factors in the Pinna illusion. *Vision Research*, *46*, 1823–1837.

Gurnsey, R., Sally, S., Potechin, C., & Mancini, S. (2002). Optimising the Pinna-Brelstaff illusion. *Perception*, *31*, 1275–1280.

Kitaoka, A., & Ashida, H. (2003). Phenomenal characteristics of the peripheral drift illusion. *Vision*, *15*(4), 261–262.

Mather, G. (2000). Integration biases in the Ouchi and other visual illusions. *Perception*, *29*, 721–727.

Morgan, M. (2002, January 24). Running rings around the brain. *The Guardian*.

Nakayama, K., & Silverman, G. (1988). The aperture problem: II. Spatial integration of velocity information along contours. *Vision Research*, *28*, 747–753.

Pinna, B. (1990). *Il Dubbio sull'apparire*. Padua, Italy: Upsel Editore.

Pinna, B. (1992). Illusioni ed illusorietà. *Contributi dei Dipartimenti e degli Istituti Italiani di Psicologia*, *5*, 221–263.

Pinna, B. (2008). The illusion of art. In B. Pinna (Ed.), *Art and perception: Towards a visual science of art*. Leiden, The Netherlands: Brill, 1–32.

Pinna, B. (2012). Illusion and illusoriness: New perceptual issues and new phenomena. In L. Albertazzi (Ed.), *The science of appearances*. New York: Blackwell-Wiley.

Pinna, B., & Brelstaff, G. J. (2000). A new visual illusion of relative motion. *Vision Research*, *40*, 2091–2096.

Pinna, B., & Gregory, R. L. (2002). Shifts of edges and deformations of patterns. *Perception*, *31*, 1503–1508.

Pinna, B., & Gregory, R. L. (2009). Continuities and discontinuities in motion perception. In G. Minati, M. Abram & E. Pessa (Eds.), *Processes of emergence of systems and systemic properties: Towards a general theory of emergence* (pp. 765–775). New York, World Scientific.

Pinna, B., & Spillmann, L. (2002). A New illusion of floating motion in depth. *Perception*, *31*, 1501–1502.

Pinna, B., & Spillmann, L. (2005). New illusions of sliding motion in depth. *Perception*, *34*, 1441–1458.

Purves, D., & Lotto, R. B. (2003). *Why we see what we do: An empirical theory of vision*. Sunderland, MA: Sinauer Associates.

Shapiro, A. G., Charles, J. P., & Shear-Heyman, M. (2005). Visual illusions based on single-field contrast asynchronies. *Journal of Vision*, *5*, 764–782.

Weiss, Y., & Fleet, D. J. (2002). Velocity likelihoods in biological and machine vision. In R. P. N., Rao, B. A. Olshausen, & M. S. Lewicki (Eds.), *Probabilistic models of the brain: Perception and neural function* (pp. 77–96). Cambridge, MA: MIT Press.

Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, *5*, 598–604.

Zöllner, F. (1860). Über eine neue art von pseudoskopie und ihre beziehungen zu den von plateau und oppel beschriebenen bewegungsphänomenen. *Annalen der Physik und Chemie*, *110*, 500–523.

Chapter 75
Color Wagon-Wheel Illusion

Arthur G. Shapiro and William Kistler

INTRODUCTION

The wagon-wheel illusion, commonly observed in cinema, consists of a simple wheel with spokes radiating from the center; the wheel physically rotates in one direction but is perceived as rotating in the opposite direction. The effect can be traced back to Faraday (1831/1859) and has recently seen a resurgance of interest (see Purves, Paydarfar, & Andrews, 1996 and VanRullen & Koch, 2003). The wagon-wheel effect depends on the aliasing of the original motion; that is, since the frame rate of filming cannot capture every moment of the wheel's rotation, the captured images provide nonsequential data of the actual rotation, and, at specific sampling rates, the perceived direction of rotation appears to be the opposite of the actual direction of rotation.

Perhaps not surprisingly, extensive and detailed studies of how and when the rotation switches from actual to reversed motion have often found that nearest-neighbor relationships between spokes more accurately predict the direction and speed of the perceived motion, rather than predicting the direction and speed of the actual, physical motion (Finlay & Dodwell, 1987; Finlay, Dodwell, & Caelli, 1984; Kline, Holcombe, & Eagleman, 2004; Kornmeier & Bach, 2012; Martineau & Glass, 2009). Whether the wagon-wheel effect exists under continuous illumination has also been debated (Andrews & Purves, 2005; VanRullen, 2006, 2007); the assumption underlying these studies is that a perceived reversal of motion under steady illumination may be indicative of the "sampling rate of consciousness."

As we show in this chapter, the standard wagon-wheel illusion can be considered an ambiguous stimulus: that is, at the sampling rates that produce the illusion, the wagon wheel contains information for motion in the actual direction of rotation as well as information for motion in the opposite direction. The visual system must (and in most cases does) decide on one of these two interpretations. Here we present a modified version of the wagon-wheel illusion that uses color to disambiguate the two directions of motion; we refer to this version as the color wagon-wheel illusion (Shapiro, Kistler, & Rose-Henig, 2012). The color wagon-wheel illusion is useful and interesting because it provides a straightforward method for separating the many different processes for representing motion in the visual system and, therefore, allows for the analysis of underlying motion mechanisms. As we demonstrate, the visual system's motion responses can act independently under some conditions but not others.

STANDARD WAGON-WHEEL EFFECT

Figure IV.75-1, an example of the standard wagon-wheel effect, consists of a ring of 12 black dots, each dot separated from the neighboring dots by 30°. Dynamic versions of Figure IV.75-1 can be seen in Interactive Figure IV.75-1 and in Videos IV.75-1a, IV.75-1b, and IV.75-1c. Toward the bottom of Interactive Figure IV.75-1, a slider allows changes to the rate of rotation (i.e., the step size of the dots, per frame).

The main observations are this: when the dots rotate slowly clockwise (e.g., 1°/frame), viewers perceive clockwise rotation; when the rate of speed is increased to 5° or 10°/frame, viewers still see clockwise motion. When the rate of rotation is increased to 15°/frame, the rotation is usually interpreted as either clockwise or counterclockwise (some observers may perceive pulsing disks rather than motion). Increasing the rotation rate to 16°/frame produces the standard wagon-wheel illusion: the perception is that of counterclockwise motion even though the dots are physically moving in a clockwise direction. The wagon-wheel illusion occurs when the ring's rotation is between 16° and 29°/frame, and, as expected, the counterclockwise rotation appears slower for the higher rotation rates in this range: that is, 29°/frame creates the perception of slow counterclockwise rotation (as if the ring is moving counterclockwise at 1°/frame) and 25°/frame creates the perception of faster counterclockwise rotation (as if the ring is moving counterclockwise at –5°/frame). It is worth noting that at 30°/frame the ring appears stationary since with every frame update, each disk moves to a location previously occupied by another disk; hence the sequence of images produced by a 30°/frame rotation is the same as the

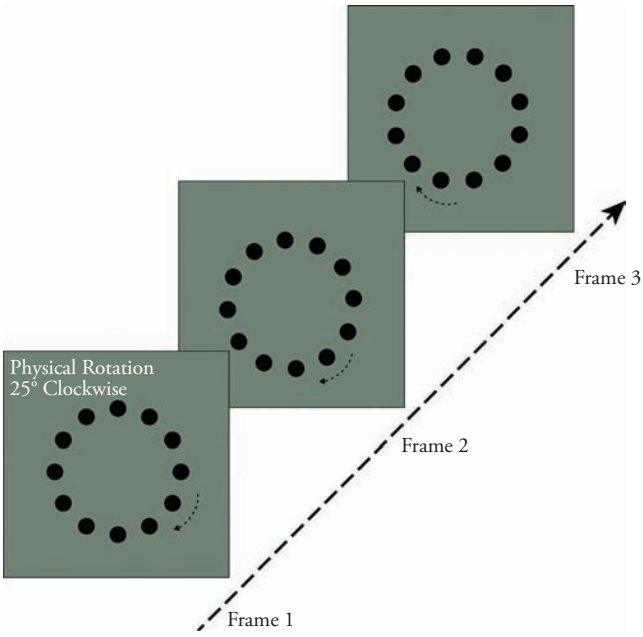


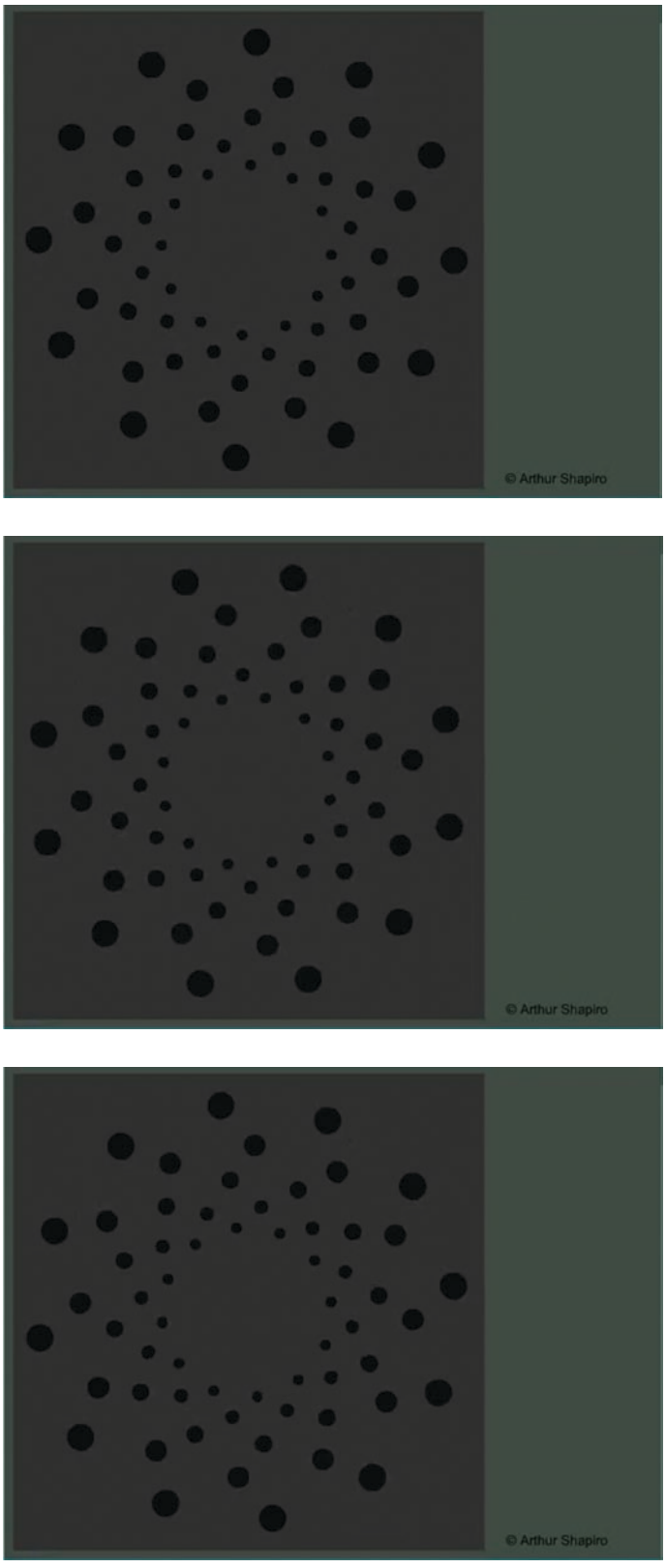
Figure IV.75-1. Twelve disks are spaced along a ring, each separated from each other by 30°. When the ring rotates clockwise with a step size of 25° per frame, the perception is that the ring rotates in the counterclockwise direction at a rate of 5° per frame. Interactive Figure IV.75-1 is an interactive version of Figure IV.75-1. Since the interactive demonstration may not function on all viewing devices, we have also created three videos that present three conditions: the rotation of the ring at 5°/frame (Video IV.75-1a; clockwise rotation perceived); 15°/frame (Video IV.75-1b; ambiguous rotation perceived); and 25°/frame (Video IV.75-1c; counterclockwise rotation perceived).

sequence of images produced when the ring of disks does not move. Between 31° and 35°/frame, however, the perception is again of clockwise motion, but each dot appears to be moving at a rate of –1° to –5°/frame.

COLOR WAGON-WHEEL EFFECT

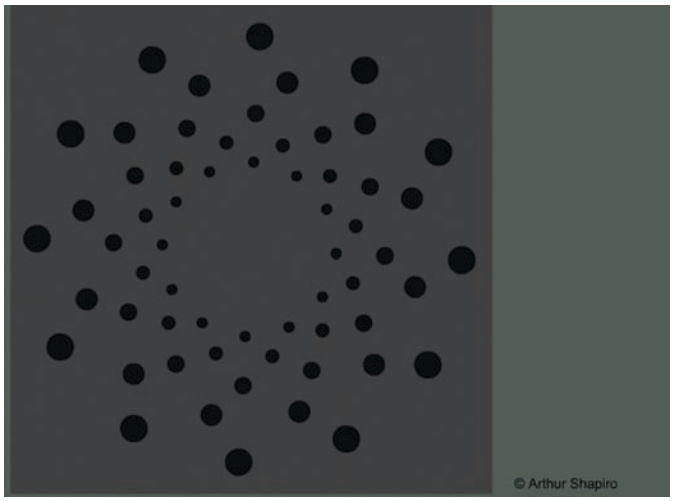
The standard wagon-wheel display is an ambiguous stimulus. The sequence of images produced by a clockwise rotation of 25°/frame is identical to the sequence of images produced by counterclockwise rotation of –5°/frame (and also identical to the sequence of images produced by rotations of 85°, 55°, –35°, –65°, –95°/frame, etc.). Therefore, since the elements of the circle are identical black dots, a 25°/frame clockwise rotation cannot be distinguished from a –5°/frame counterclockwise rotation. We perceive counterclockwise rotation presumably because the visual system interprets the information from motion that resembles repeated shifts of –5°/frame as much stronger (or as much more likely) than the motion information from the actual repeated shifts of 25°/frame or any other interpretation.

The ambiguity found in the standard wagon-wheel illusion can be overcome by changing the appearance of one or more of the dots so that they are no longer identical. For instance, in Figure IV.75-2 we change one of the 12 black dots from Figure IV.75-1 to a yellow dot. One might hypothesize that the addition of the yellow dot would make the reversal



Video IV.75-1. See caption for Figure IV.75-1.

of motion in the wagon-wheel illusion disappear; after all, if the visual system sees the whole ring rather than its parts, the yellow dot would indicate that the whole ring moves clockwise when set to rotate at 25°/frame. Instead,



Interactive Figure IV.75-1. See caption for Figure IV.75-1.

we perceive two types of motion, both tied to the local information in the display: the black dots appear to move counterclockwise (as in the standard wagon-wheel illusion in Fig. IV.75-1), while the yellow dot appears to move and is physically moving clockwise.

Dynamic versions of Figure IV.75-2 can be seen in Interactive Figure IV.75-2 and in Videos IV.75-2a, IV.75-2b, and IV.75-2c. Interactive Figure IV.75-2 allows the observer to adjust the rate of rotation. For the black dots, changing the step size produces the same effect as in Interactive Figure IV.75-1, but for any adjustment of step size, the yellow dot always appears to rotate clockwise. It is

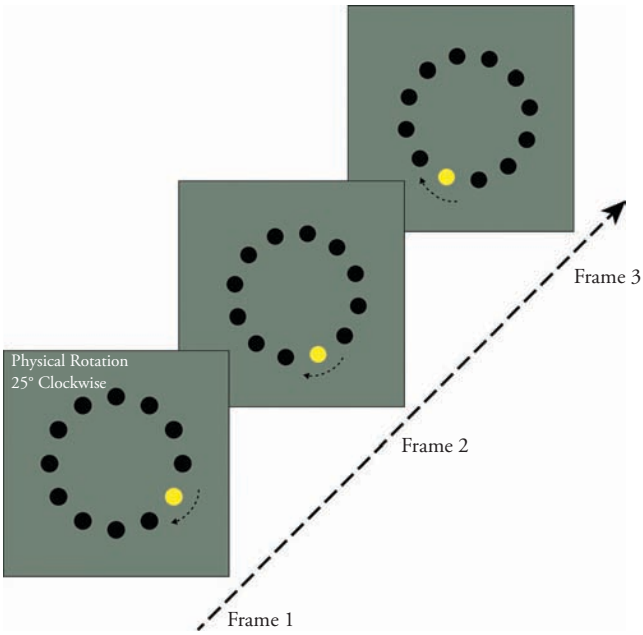
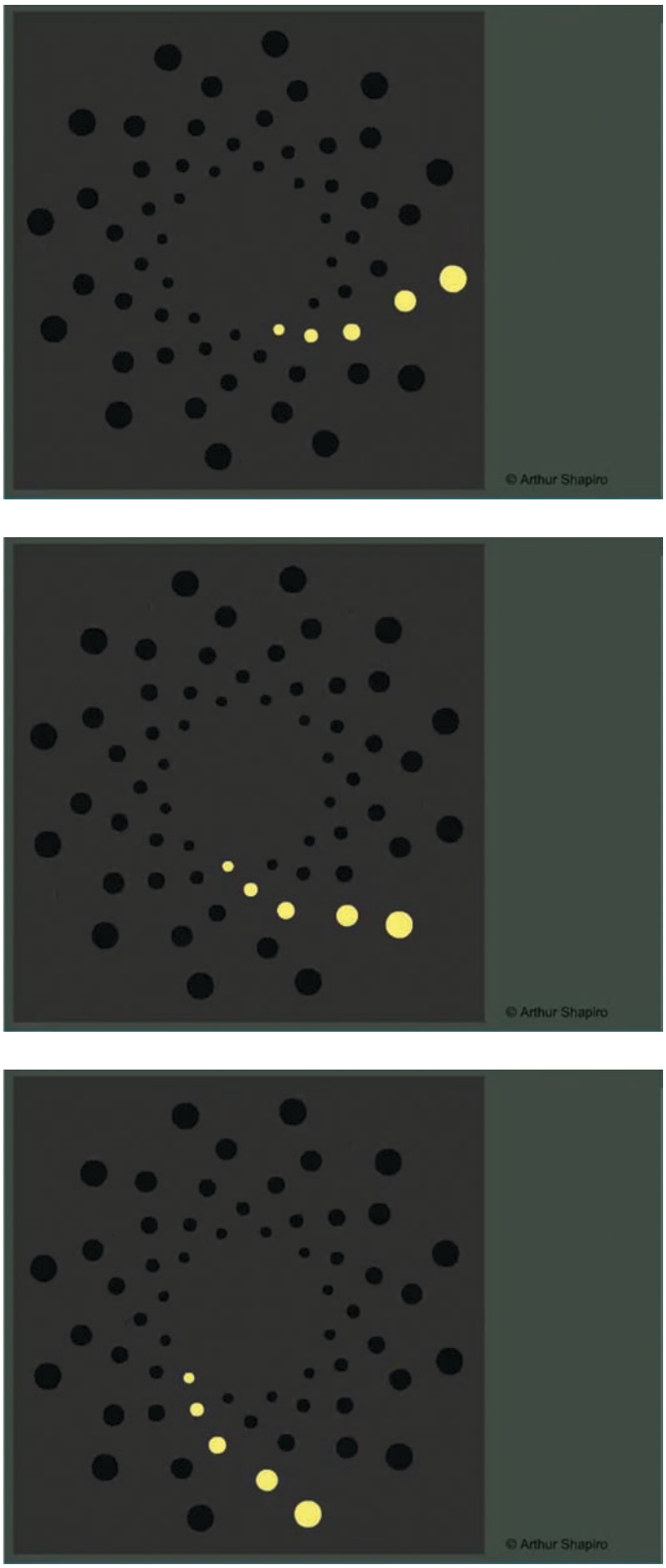
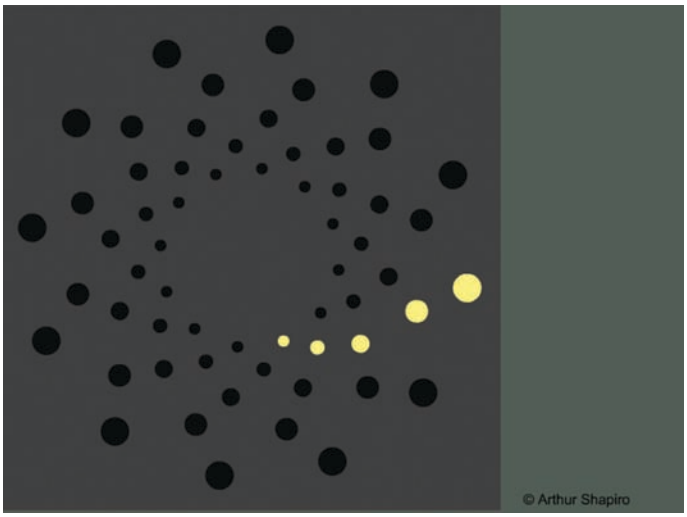


Figure IV.75-2. When the ring rotates clockwise at 25°/frame, the ring appears to rotate in the counterclockwise direction at 5°/frame, but the yellow disk appears to rotate clockwise. Two opposing directions of motion are therefore perceived. Interactive Figure IV.75-2 provides an interactive version of Figure IV.75-2. Videos IV.75-2a, IV.75-2b, and IV.75-2c show the appearances of the ring at three rotation speeds: 5°/frame (Video IV.75-2a; black and yellow dots perceived as moving clockwise); 15°/frame (Video IV.75-2b; ambiguous rotation of black dots and clockwise motion of yellow dot); and 25°/frame (Video IV.75-2c; black dots perceived as moving counterclockwise and yellow dot perceived as moving clockwise.)



Video IV.75-2. See caption for Figure IV.75-2.



Interactive Figure IV.75-2. See caption for Figure IV.75-2.

particularly illuminating to set the step size to 15°/frame, since the movement of the 12 black dots in Figure IV.75-1 at this step size presents the ambiguous perception of clockwise or counterclockwise motion. Yet even in this ambiguous condition, in which the black dots have indeterminate direction, the yellow dot in Figure IV.75-2 is perceived as rotating clockwise. At 30°/frame, the black dots appear to remain still, while the yellow dot continues to appear to move clockwise along the ring.

EQUILUMINANT WAGON-WHEEL EFFECT

The color wagon-wheel illusion in Figure IV.75-2 and Video IV.75-2 creates the simultaneous perception of two opposing directions of motion. A number of researchers have proposed that the visual system contains several ways of constructing motion; one currently popular taxonomy, the three-system functional architecture of Lu and Sperling

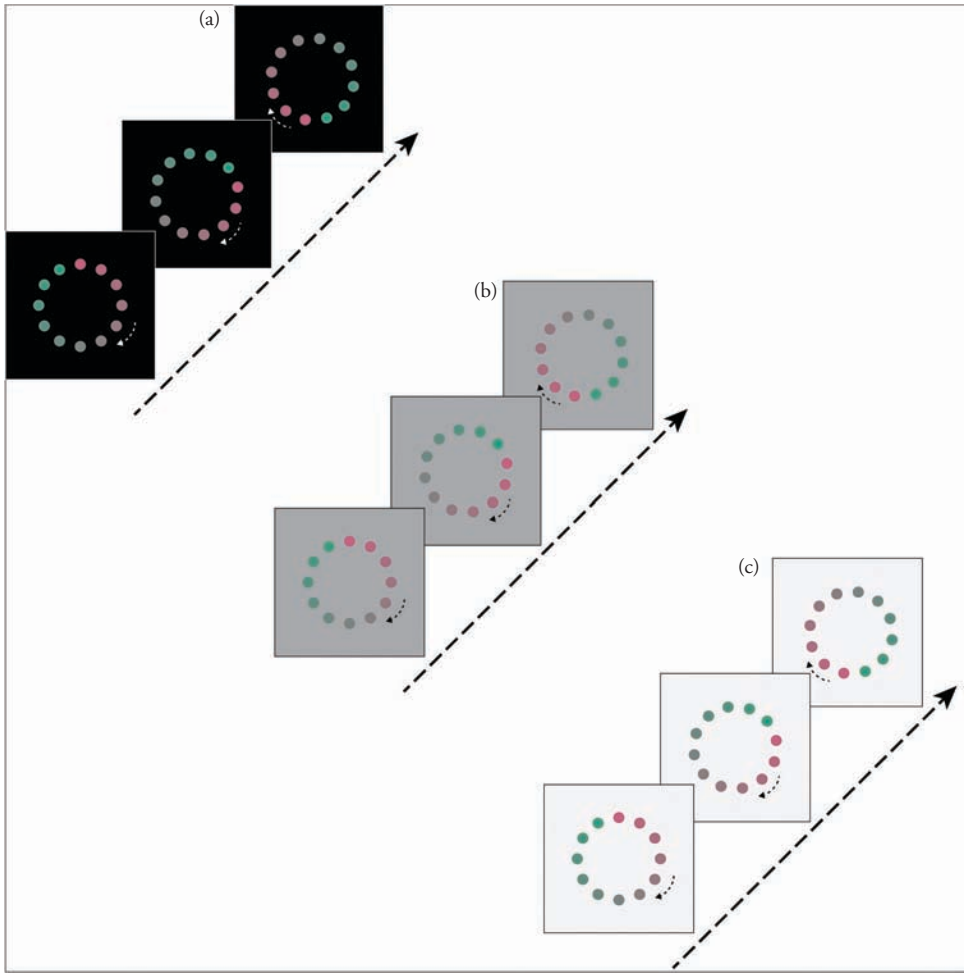
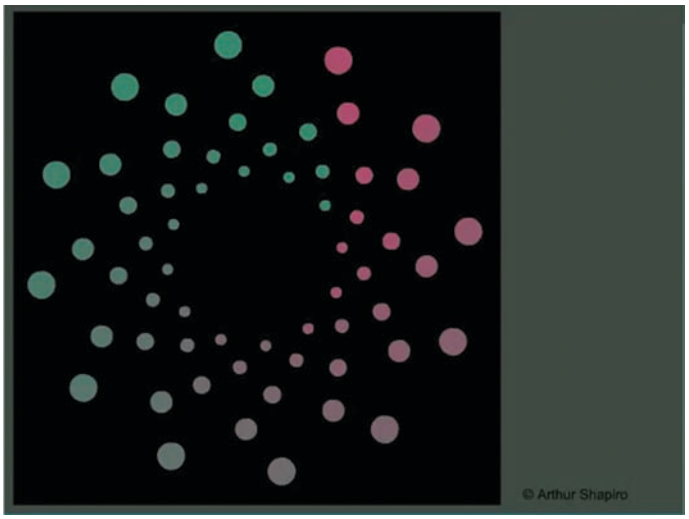
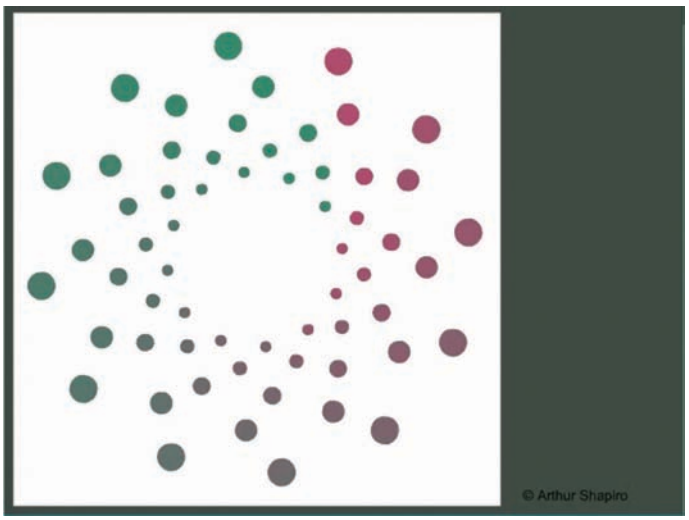
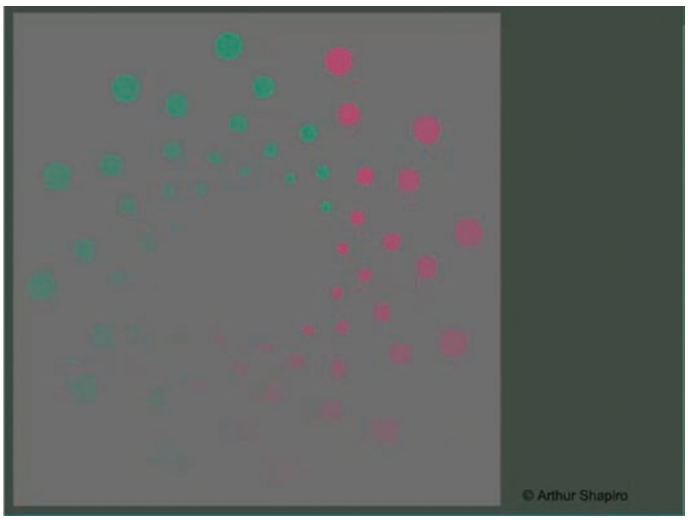


Figure IV.75-3. Twelve dots are rendered with varying hues but identical luminance values. The background luminance of panels a, b, and c varies from dark to light (pixel values zero, 128, 255). In (a) and (c), when the ring rotates clockwise at 25°/frame, the ring appears to rotate counterclockwise, but the hues of the ring appear to rotate clockwise when the observer attends to a single colored dot. In panel (b), the perception of the counterclockwise motion nearly disappears, and the perception is dominated by the clockwise motion. Interactive Figure IV.75-3 allows the observer to adjust the speed of rotation and the luminance of the background. Videos IV.75-3a, IV.75-3b, and IV.75-3c show the appearances of the ring at the three levels of background luminance depicted in panels (a), (b), and (c), all with a rotation speed of 25°/frame.







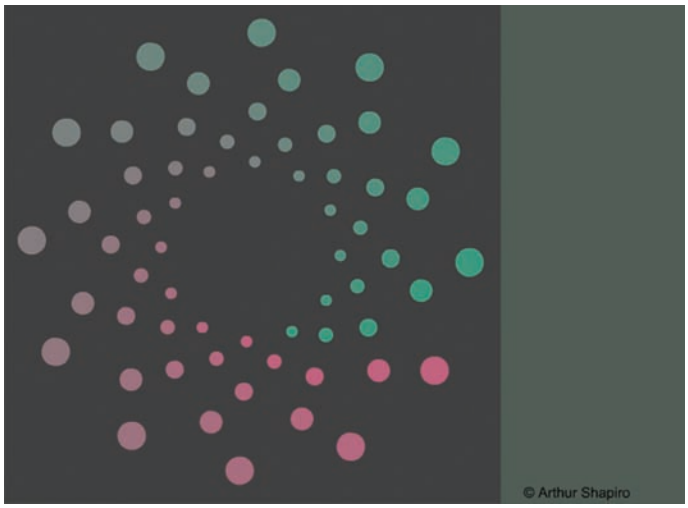
Video IV.75-3. See caption for Figure IV.75-3.

(2001), can be summarized as a *first-order system* that extracts motion from moving luminance modulation using a motion-energy algorithm (or, equivalently, a Reichardt detector), a *second-order system* that extracts motion from contrast deviations in surrounding fields, and a *third-order system* based on being able to track the motion of an element or object in a stimulus. The question is whether the two directions of motion visible in Interactive Figure IV.75-2 (at the noted rates of rotation) can be used to determine whether two types of motion systems are responsible for the perception of motion in two directions.

In principle, the perception of clockwise and counterclockwise motion could result from one or more of the first-, second-, and third-order motion systems. However, since first- and second-order systems respond primarily to changes in luminance, we can try to isolate responses to third-order motion by removing luminance contrast from the display. To this end, we created a stimulus in which all 12 dots have different hues but are close to the same luminance. We note that the dots are close to equiluminant instead of accurately equiluminant because equiluminance is difficult to establish: the wagon-wheel illusion can be viewed on different types of devices with different luminance settings, and people obviously differ in their luminance response to chromatic lights.

Figure IV.75-3, Interactive Figure IV.75-3, and Videos IV.75-3a, IV.75-3b, and IV.75-3c show the 25°/frame equiluminant wagon wheel. Interactive Figure IV.75-3 allows adjustments to the background luminance: when the background is darker or brighter than the dots (as in Videos IV.75-3a and IV.75-3c), the dots appear to move counterclockwise, yet an observer who attends to the colors of the dots can also track the dots moving clockwise. However, when the background and the dots in Interactive Figure IV.75-3 are equiluminant (as in Video IV.75-3b), the strength of the counterclockwise motion is diminished, and the dots seem to move primarily in a clockwise direction.

This finding has several implications: (a) the motion signal that produces the counterclockwise perception is likely a first-order motion signal (or another motion system that is weaker at equiluminance); (b) the perception of counterclockwise motion is not a function of the physical motion of the dots but is rather a function of the visual system calculating the motion from the contrast between the dots and the surrounding field; and (c) the perception of clockwise motion is the result of a motion system that is



Interactive Figure IV.75-3. See caption for Figure IV.75-3.

not affected by equiluminance, presumably due instead to a feature-tracking or third-order motion system.

EQUILUMINANT WAGON-WHEEL EFFECT WITH INTERNAL MOTION

When the dots and the background in Interactive Figure IV.75-3 are close to equiluminant (as in Video IV.75-3b), motion is perceived primarily in the clockwise direction, but some motion is visible in the counterclockwise direction. Since the perception of counterclockwise motion is driven by first- or second-order motion systems, and since these systems respond to luminance, we could presumably eliminate the perception of counterclockwise motion by asking observers to adjust the colored lights so that they match each other in luminance. Yet, as mentioned earlier, such adjustments are not always easy to make and limit the generality of the phenomenon. Another approach is to decrease the response of the first- and second-order systems relative to the motion that arises from tracking the equiluminant colors.

To this end, we created the elements in Figure IV.75-4, Interactive Figure IV.75-4, and Videos IV.75-4a, IV.75-4b, and IV.75-4c. Instead of dots, the elements are rectangular bars that individually rotate clockwise around their own axis and, additionally, rotate as a ring of bars clockwise around its own center point. Interactive Figure IV.75-4 allows adjustments to the rotation rate, the background luminance, and the color conditions of the bars in order to replicate the conditions of Figures IV.75-1 and IV.75-2. Videos IV.75-4a, IV.75-4b, and IV.75-4c have a step size of 25°/frame. Videos IV.75-4a and IV.75-4c show the equiluminant bars set against a background of contrasting color (in Video IV.75-4a the background is black, and in Video IV.75-4c the background is white). The dual motion of the noncircular wheel elements makes it more difficult for the visual system to track global motion in the counterclockwise direction.

Figure IV.75-4 and Video IV.4a and IV.4c initially result in perceptions like those associated with the equiluminant wagon-wheel illusion in Videos IV.75-3a and IV.75-3c: against a black (IV.75-4a) or white (IV.75-4c) background, an observer who is not focusing on an individual element perceives counterclockwise rotation. However, if the observer focuses on a particular color in this version of the color wagon-wheel effect (e.g., green), all of the elements of the wheel appear to move clockwise—the actual direction of motion. When the background luminance is matched to the luminance of the bars (as in Video IV.75-4b), the wagon-wheel effect completely disappears without the need to focus on an individual wheel element: the

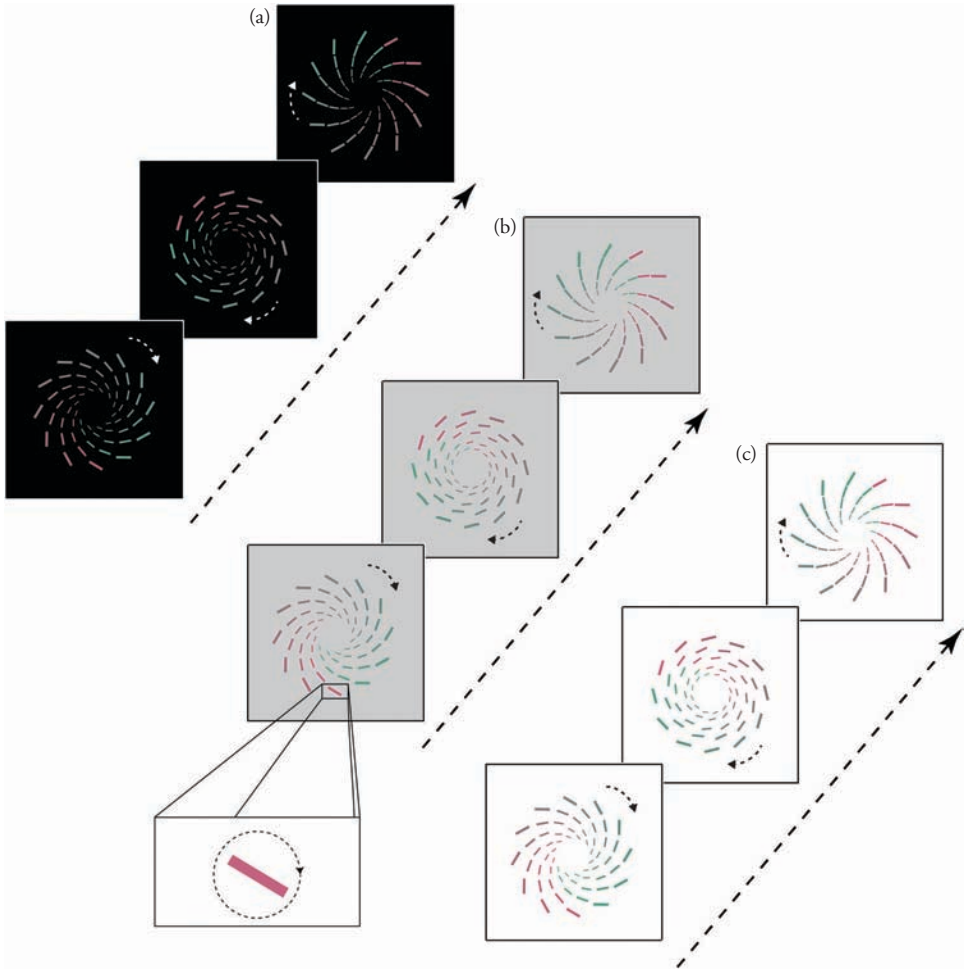
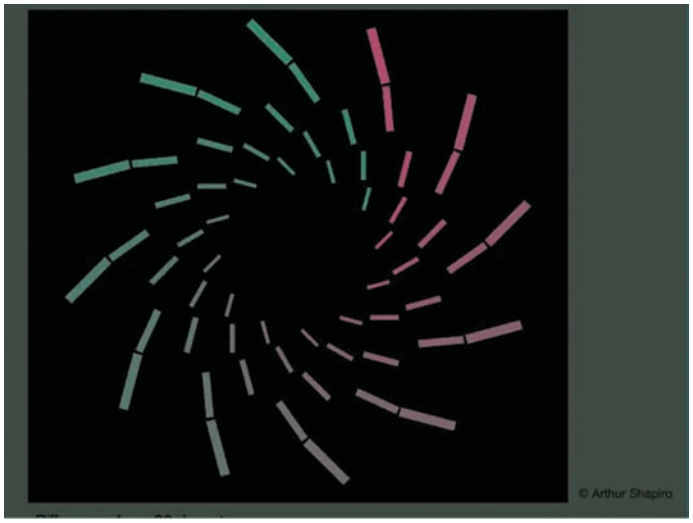


Figure IV.75-4. Similar to Figure IV.75-3, except that the dots are replaced with rectangular bars that rotate not only in a ring but also individually, each on its own axis. See also Interactive Figure IV.75-4 and Videos IV.75-4a, IV.75-4b, and IV.75-4c.



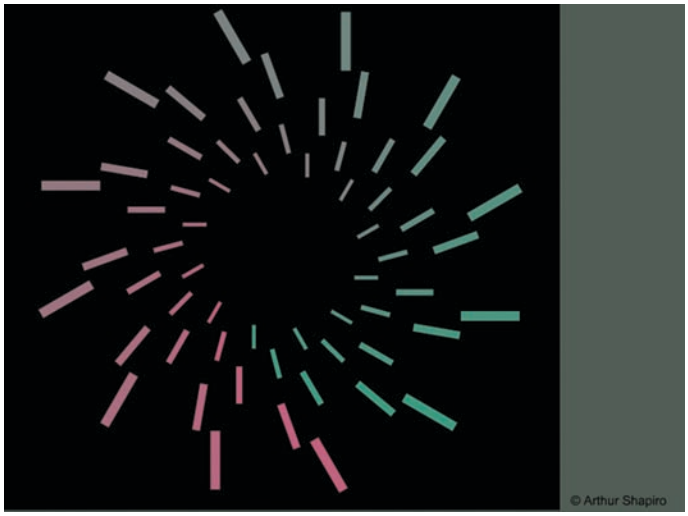




Video IV.75-4. See caption for Figure IV.75-4.

perception of the ring's rotation is in accord with the actual clockwise rotation of the colored bars.

The rotation of equiluminant bars around their own axes therefore seems to reduce the strength of the first- and second-order motion responses relative to the strength



Interactive Figure IV.75-4. See caption for Figure IV.75-4.

of the third-order motion response. Curiously, however, in Video IV.75-4b, if an observer fixates on a single bar, that bar seems to be moving around the ring in a counterclockwise direction—that is, the effect is the inverse of the equiluminant wagon wheel in Video IV.75-3b. Attention to a particular element seems to undermine the perception of attention-driven third-order motion in favor of a first- and/or second-order response to the individual element. We have not yet proposed an explanation for the strengthening of reverse motion under these conditions. Our experiments continue.

CONCLUSIONS

The color wagon-wheel illusion is a straightforward and useful variant of the classic wagon-wheel illusion. The classic wagon-wheel illusion is often discussed as an example of spatial-temporal aliasing; however, such explanations typically gloss over the fact that the wagon-wheel illusion contains within it an ambiguous mixture of many different types of motion information. The reverse motion that is most often perceived is one of many possible perceptual alternatives. The color wagon-wheel illusion strengthens the perception of one of the alternatives by changing the identity of one (or more) of the elements.

We have shown that the variations of the color wagon-wheel illusion can be used to separate different classes of motion phenomena. The color wagon-wheel illusion highlights that the wagon-wheel illusion (with or without differing color in one or more elements) contains more than a single source of motion information. Alternative explanations that account for these sources of motion information may be worth consideration when forming hypotheses about perceived reversals of motion.

REFERENCES

Andrews, T., & Purves, D. (2005). The wagon-wheel illusion in continuous light. *Trends in Cognitive Sciences*, *9*(6), 261–263.

Faraday, M. (1859). On a peculiar class of optical deceptions. In M. Faraday, *Experimental researches in chemistry and physics* (p. 291). London: Taylor & Francis. (Original work published 1831)

Finlay, D., Dodwell, P., & Caelli, T. (1984). The wagon wheel effect. *Perception*, *13*(3), 237–248.

Finlay, D., & Dodwell, P. (1987). Speed of apparent motion and the wagon wheel effect. *Perception & Psychophysics*, *41*(1), 29–34.

Kline, K., Holcombe, A., & Eagleman, D. M. (2004). Illusory motion reversal is caused by rivalry, not by perceptual snapshots of the visual field. *Vision Research*, *44*, 2653–2658.

Kornmeier, J., & Bach, M. (2012). Ambiguous figures—What happens in the brain when perception changes but not the stimulus. *Frontiers in Human Neuroscience*, *6*(5), 1–

Lu, Z., & Sperling, G. (2001). Three-systems theory of human visual motion perception: Review and update. *Journal of the Optical Society of America A*, *18*(9), 2331–2370.

Martineau, P., Aguilar, M., & Glass, L. (2009). Predicting perception of the wagon wheel illusion. *Physical Review Letters*, *103*(2), 028701.

Nishida, S. (2011). Advancement of motion psychophysics: Review 2001–2010. *Journal of Vision*, *11*(5), 1–53.

Purves, D., Paydarfar, J. A., & Andrews, T. J. (1996). The wagon wheel illusion in movies and reality. *Proceedings of the National Academy of Sciences USA*, *93*, 3693–3697.

Shapiro, A. G., Kistler, W., & Rose-Henig, A. (2012). Color wagon-wheel illusion. Winner: Best illusion of the year contest. http://illusionoftheyear.com/2012/color-wagon-wheel/

VanRullen, R. (2005). Attention-driven discrete sampling of motion perception. *Proceedings of the National Academy of Sciences USA*, *102*, 5291–5296.

VanRullen, R. (2006). The continuous wagon wheel illusion is object-based. *Vision Research*, *46*(24), 4091–4095.

VanRullen, R. (2007). The continuous wagon wheel illusion depends on, but is not identical to neuronal adaptation. *Vision Research*, *47*(16), 2143–2149.

VanRullen, R., & Koch, C. (2003). Is perception discrete or continuous? *Trends in Cognitive Science*, *7*, 207–213.

Chapter 76
The Aperture Problem

Illusions Arising During the Integration and Segmentation of Motion Within and Across Apertures

Maggie Shiffrar

ILLUSORY MOTION PERCEPTS WITHIN AND ACROSS APERTURES

What Is the Aperture Problem?

As I look outside my office window, I see a man walking his dog, a car rolling by, and shrubbery swaying in the breeze. To interpret such a complex array of motions accurately, my visual system must simultaneously perform two apparently conflicting tasks. On one hand, it must separate motion signals belonging to different objects. It would be an error to combine the motion of the car with the movements of the dog. Concurrently, my visual system must also integrate motion signals belonging to the same object. While the arms of a walking person typically swing in opposition, their motion signals must nonetheless be combined before the person's movements can be accurately interpreted. These simultaneous processes of integration and segmentation are what allow us to perceive static and moving objects in our world. Indeed, classic Gestalt grouping principles essentially prescribe which portions of any image are seen as belonging together (integrated) and which portions are perceptually separated (segmented). Since our visual perception of motion seems effortless and typically unambiguous, it would be easy to suppose that the processes of motion segmentation and motion integration are simple and straightforward. Yet, as will become increasingly clear, the accurate interpretation of visual motion signals is a challenging problem.

Let us start with the question of why: Why is it such a challenge to determine which motion signals should be grouped together and which motion signals should be separated? Much of the difficulty arises from the fact that the inputs to the motion perception system (i.e., initial motion measurements) are inherently ambiguous. This ambiguity in turn arises from fundamental characteristics of the neurons that perform early measurements of visual image motion (e.g., Hubel & Wiesel, 1968; Movshon, Thompson, & Tolhurst, 1978). This chapter examines two of these interrelated characteristics—namely, conjunctive sensitivity and receptive field size—and describes some of the conditions under which motion ambiguity gives rise to illusions.

First, early receptors that are responsive to motion are conjointly sensitive to direction and orientation. As one illustration of this issue, consider the Nobel Prize–winning research of David Hubel and Torsten Wiesel (e.g., 1959, 1968) in which they discovered the receptive field properties of cortical neurons in early processing stages of the visual cortex. This groundbreaking work uncovered neurons in the striate cortex that respond vigorously when a stationary or translating luminance edge of a particular orientation is projected within their receptive field (see Hubel, 1982, for review). The key point for the current discussion is that in these groundbreaking studies, the direction of edge translation and edge orientation were always confounded. Specifically, the luminance-defined edge always translated in a direction perpendicular to its orientation. And there is a very good reason for this confound; namely, these neurons can only detection motion that is perpendicular to the orientation of a translating edge. Indeed, early motion sensitive cells are individually blind to all components of motion that are parallel to the orientation of a straight edge translating within their receptive fields. As a result, different combinations of stimulus orientation and direction can give rise to the same neuronal response. This ambiguity is known as the aperture problem (Stumpf, 1911; Todorović, 1996; Wallach, 1935; Wuerger, Shapley, & Rubin, 1996) and is illustrated in Figure IV.76-1.

To understand the aperture problem, first consider that the motion of any line can be decomposed into the portion that is parallel to the line and the portion that is perpendicular to the line. Whenever the ends of a line fall outside of a neuron's receptive field, that neuron cannot detect, or "see," any of the motion that is parallel to the line's orientation (i.e., along the length of the line). As a result, a neuron can only detect the perpendicular component of the line's motion (Fig. IV.76-1). Because only this perpendicular component of motion can be measured, all motions having the same perpendicular motion will appear to be identical even if these motions differ significantly in their parallel component. For this reason, a neuron will give the same response

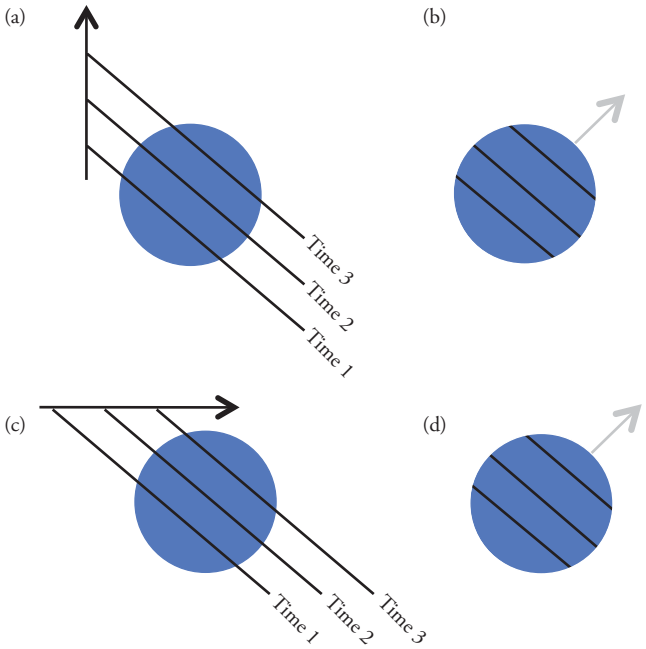


Figure IV.76-1. The aperture problem. (a) A line translating vertically is seen at three different time points. The line's ends extend beyond the blue receptive field. Because the ends of the lines cannot be detected within the receptive field, the component of motion parallel to the line's orientation cannot be detected within that receptive field. As a result, (b) only the component of motion perpendicular to the orientation of the line can be detected. In other words, vertical motion within this aperture gives rise to the illusory perception of diagonal motion. (c) A horizontally translating line is seen through the same blue receptive field. Again, because the ends of the lines cannot be detected within the receptive field, (d) only the component of motion perpendicular to the line's orientation can be detected. As a result, while the horizontal and vertical directions of translation (a) are clearly different, the receptive field cannot detect this difference (b, c, and d). In the neural domain, motion-sensitive neurons with small receptive fields commonly show the same response to very different motions. Thus the end result of the aperture problem is that early motion measurements are often ambiguous. (Figure by Maggie Shiffrar.)

to many different motions. Because all known visual systems, whether biological or computational, have receptors that are limited in size, this measurement ambiguity has been extensively studied (e.g., Hildreth, 1984; Pack, Gartland, & Born, 2004; Stumpf, 1911; Wallach, 1935).

A second and not unrelated reason why initial motion measurements are inherently ambiguous is because neurons in the early stages of the visual system have tiny receptive fields and as such can only respond to changes within very small regions of an observer's field of view. As a result, when the image of a real moving object is projected across an observer's retina, each motion-sensitive receptor can only detect the motion of a tiny subregion of the object. This creates a problem for the visual system that is analogous to John Godfrey Saxe's (1855) rendition of the classic Indian tale of six blind men feeling an elephant in order to identify it. One man feels only the elephant's tusk. Another man feels a tail. Yet another feels an ear, and so on. Individually, no man can identify the elephant by generalizing from the tactile information gathered from one small region of the elephant's body. Instead, the identification of the elephant requires a complex, higher-level integration of the men's collective experiences. In the same way, no individual neuron in an observer's visual system can identify an object's motion. Instead, visual percepts of object motion require complex, higher-level processes of integration, and separation, across both space and time.

The end result of both of these characteristics of early motion receptors is that the interpretation of visual motion requires the integration and segmentation of an enormous matrix of individually ambiguous velocity estimates. How does the visual system solve this complex problem?

The Barber Pole Illusion: Motion Integration Within an Aperture

For over 75 years (e.g., Wallach, 1935), several researchers have used the barber pole illusion to study how our visual systems might resolve the ambiguity of motion signals within a single area or receptive field. The barber pole illusion originally came from, not surprisingly, the rotating striped poles (see Fig. IV.76-2a) located outside barbershops. Since the Middle Ages, these traditionally red-and-white striped poles have announced the services of barbers. As an aside, for centuries barbers performed significant medical procedures, including surgery and bloodletting, in addition to cutting hair (MacNulty, 1945). Indeed, the red and white stripes of a traditional barber pole represent the bloodied and clean bandages associated with blood-letting.

What is the illusory motion to which rotating barber poles give rise? As illustrated in Figure IV.76-2a, because the pole rotates about a vertical axis, the physical motion of the stripes is necessarily and entirely horizontal. However, when one watches a barber pole, the stripes appear to translate vertically. In other words, physically horizontal motion gives rise to the perception of vertical motion. How does this illusory percept occur? The key contributor is the vertical orientation of the underlying cylinder. First, for purposes of simplification, let us flatten the visible portion of the barber pole cylinder to a flat surface (Fig. IV.76-2b). Recall that early measurements of motion are made through small receptive fields, or apertures, that produce ambiguous measurements of translating edges. Let us assume that each stripe of the barber pole is measured by just three apertures: one aperture on the ends of each stripe and one covering the middle of each stripe (top line of Fig. IV.76-2b). The center aperture "sees" only the homogeneous portion of the stripe, and, in concordance with the aperture problem, cannot detect the component of translation that is parallel to the stripe's orientation. As a result, the center aperture produces ambiguous motion estimates (cluster of gray arrows). The two outside apertures, however, fall on the ends of the stripe. Both the parallel and the perpendicular components of translation can be easily measured from two-dimensional discontinuities such as corners and ends. Thus unambiguously vertical motion (black arrows) can be detected in the two outside apertures. In this barber pole scenario, the same combination of unambiguously vertical motion would be detected on the ends of each stripe, and ambiguous motion would be detected in the center portion of each stripe (Fig. IV.76-2b). Assuming that the visual

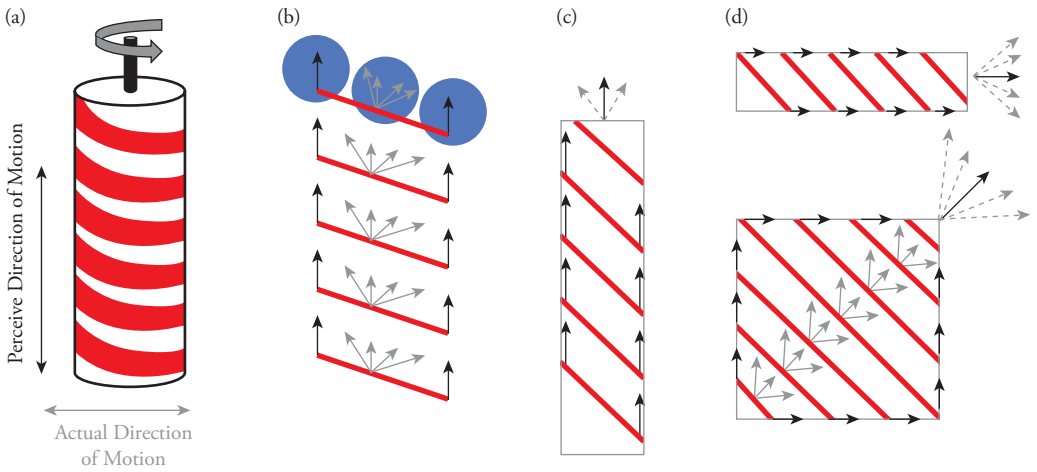


Figure IV.76-2. The barber pole illusion. (a) In a traditional barber pole, a stripped cylinder rotates about a central axis. While the physical motion of the stripes is horizontal (gray arrow), observers perceive vertical motion (black arrow). (b) A projection of a barber pole display onto a flat, two-dimensional surface. The three blue circles represent three receptive fields through which the velocity of each of the horizontally translating stripes is measured. The visible endpoints of each stripe (created at the point of contact with the viewing window) appear to translate vertically (black arrows). Because of the aperture problem, the motion of the central portion of each stripe is ambiguous (gray arrows). The visual system relies on the vertical motions of the stripes' endpoints rather than on the ambiguous velocity estimates from the central portion of each stripe. As a result, the stripes in the barber pole illusion appear to move vertically. (c) While the stripes move horizontally in traditional barber poles, the same vertical motion (large black arrow at tope) is perceived when the stripes move in other directions (dashed gray arrows) because the visual system continues to rely on the vertical motion of the line endings. (d) The shape of the barber pole (or aperture) determines the direction in which the stripes appear to translate. If a rectangular window is oriented horizontally, then the stripes will appear to translate horizontally because most of the visible endpoints appear to translate horizontally. If the stripes translate behind a square aperture, then the conflicting motions of the endpoints will cancel each other out and the stripes will appear to translate in the direction perpendicular to their orientation, just as in the case of a circular aperture (Fig. IV.76-1). (Figure by Maggie Shiffrar.)

system takes advantage of the degree of certainty associated with different types of motion measurements (more on this later), it is not surprising that the visual system would rely heavily on the unambiguous vertical measurements and as a result produce the visual perception of vertical motion that defines the barber pole illusion. In this example, the physical motion is always horizontal. However, the same vertical motion percepts are produced when the stripes move in other directions (gray dashed arrows in Fig. IV.76-2c).

If, while holding stripe orientation constant, the orientation of a barber pole determines the direction in which the stripes appear to translate, then changes in a barber pole's (or aperture's) orientation should produce changes in the perceived direction of stripe translation. In Figure IV.76-2d, the barber pole is oriented horizontally. Under this condition, the same diagonal stripes now appear to translate horizontally because the ends of the stripes appear to translate horizontally. If the shape of the barber pole were changed to a square, then the ends of the stripes would produce conflicting directional cues (small black arrows at aperture edges). Such conflict increases uncertainty, which seems to lessen the influence of these signals on the overall interpretation of the strip's motion. As a result, perceived stripe translation reverts back to the direction orthogonal to stripe orientation (as with circular apertures; Fig. IV.76-1).

While the interpretation of a single translating line is ambiguous when seen through a relatively small aperture, the possible interpretations of its motion are limited to a large family of related motions. All of the members of this family differ only in the component of translation that is parallel to the line's orientation. Members of two hypothetical families are illustrated by the groups of three arrows in Figure IV.76-3a–b. Each of the gray arrows depicts one possible velocity. Arrow orientation indicates the line's potential direction of translation while the arrow length indicates the corresponding speed of translation. Notice that the arrows all line up along a dashed line. This dashed line, known as the constraint line, represents the entire family of velocities that is consistent with the motion measured from a single translating line or grating of a particular orientation (Adelson & Movshon, 1982).

The visual system can solve the aperture problem in part by combining the individually ambiguous motion estimates from two differently oriented lines or gratings, as

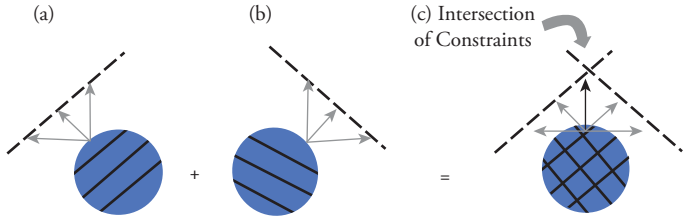


Figure IV.76-3. (a) and (b) While the motion of a line or grating viewed through a circular aperture is ambiguous, the family of motions that could have produced that motion is restricted to those that fall along the dotted line. (c) When two gratings are rigidly superimposed, then the point at which the two dotted or constraint lines intersect defines the direction in which the two gratings appear to translate (as long as those gratings are defined by similar spatial frequencies and luminance contrasts; Adelson & Movshon, 1982). (Figure by Maggie Shiffrar adapted from Adelson & Movshon, 1982.)

illustrated in Figure IV.76-3c. As long as two differently oriented lines are rigidly connected to each other, and therefore actually moving in the same direction, their corresponding constraint lines will intersect at a single point. This point, known as the intersection of constraints (IOC), defines the only possible motion interpretation that is shared by the two rigidly connected translating lines. Thus when the visual system is correct in assuming that two lines are rigidly connected to each other, then the motions of those lines can be uniquely interpreted. Experimental support for this IOC approach comes from studies examining the visual perception of and neural response to spatially overlapping edges and gratings (e.g., Adelson & Movshon, 1982 & Movshon, Adelson, Gizzi, & Newsome, 1985). The IOC approach provides one example of how the visual system might solve the aperture problem for superimposed and rigidly connected gratings presented within a single receptive field or region of visual space.

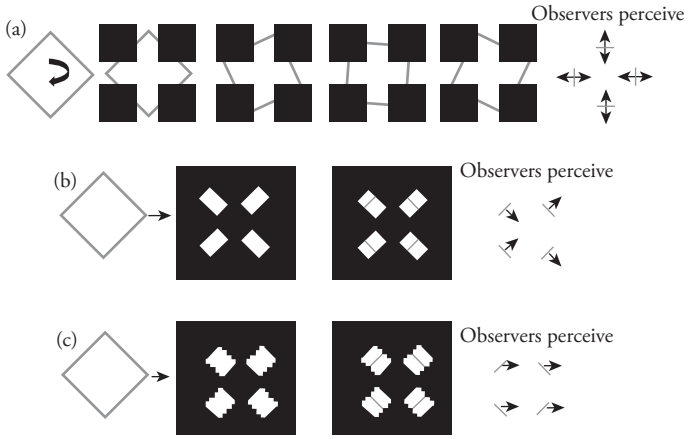


Figure IV.76-4. (a) A pulsing diamond illusion. When a diamond (gray) rotates behind a set of occluders (black), the rigid diamond appears to pulse nonrigidly in size as it rotates whenever its corners are hidden. This illusory perception results from a difficulty in integrating motion signals across space. (b) When a rigid diamond (gray) translates horizontally behind a set of apertures (black) that occlude the diamond's corners, then the diamond's visible edges appear to translate nonrigidly relative to one another. More specifically, each edge appears to translate independently in the direction perpendicular to its own orientation. (c) When the shape of the apertures, or other factors, adds noise to or increases uncertainty about the momentary positions of the ends of the diamond's visible edges, then the diamond's edges appear to translate coherently and horizontally. (Figure by Maggie Shiffrar.)

Illusory Motion Percepts Across Multiple Apertures

Whenever a visual scene contains multiple objects, as in the example that opened this chapter, then the application of the IOC or similar algorithm is insufficient to generate an accurate interpretation of the motions in the image (e.g., Majaj, Carandini, & Movshon, 2007). To understand how the visual system might analyze the motion information across different spatial locations, researchers have examined the visual perception of simple objects moving behind multiple apertures.

When observers view differently oriented edges through nonoverlapping apertures, they experience systematic difficulties in their ability to link motion signals across those disconnected edges (e.g., Lorenceau & Shiffrar, 1992; Shiffrar & Pavel, 1991). In a compelling illustration of this, a simple diamond figure is presented as rotating behind a set of stationary occluders, as illustrated in Figure IV.76-4a (Shiffrar & Pavel, 1991). Surprisingly, the rigid diamond appears to rotate normally only when its corners are visible. Whenever its corners are occluded, and thus only straight edges are visible (an "aperture problem" situation), then the diamond appears to nonrigidly expand and contract as it rotates (Shiffrar & Pavel, 1991). Even though observers know that they are viewing a diamond rigidly rotating behind four stationary apertures, they cannot accurately combine the motion signals across the straight edges of the diamond. As a result, a physically rigid diamond appears to be nonrigid as it rotates. This finding is inconsistent with classic theories that propose that the visual system overcomes the ambiguity of motion measurements by selecting image interpretations that are consistent with rigid objects (e.g., Ullman, 1979).

Does the inability to accurately integrate velocity estimates across a rigidly rotating polygon depend on the fact that rotational motion is in play? The answer is clearly no. When the same diamond figure translates horizontally behind a set of apertures, as shown in Figure IV.76-4b, each of its visible edges appears to translate diagonally in the direction perpendicular to its own orientation (Lorenceau & Shiffrar, 1992). That is, while the diamond figure translates coherently along the horizontal axis, each of its four edges appears to translate incoherently along diagonal axes. In an experimental measure of this illusion, naïve observers viewed a diamond figure translating in a clockwise or counterclockwise direction while maintaining the same upright orientation (like a car on a Ferris wheel). The diamond was viewed through a set of rectangular apertures as depicted in Figure IV.76-4b. Observers performed a direction discrimination task constructed so that the diamond's direction of translation could only be determined from an integration of the motion measurements across the diamond's visible edges. When the translating diamond was presented at high luminance contrast relative to the background, observers performed at chance levels in the direction discrimination task. That is, even though observers knew that they were viewing a rigidly translating diamond, they could not link motion signals across the diamond's sides and determine its direction of translation. Instead, under these conditions, observers perceived the diamond's apparently coherent motion as four independently translating object fragments.

The illusions described here suggest that the integration of motion signals within the same spatial region (Fig. IV.76-3) differs from the integration of motion signals across disconnected spatial regions (Fig. IV.76-4). So how does the visual system integrate motion information across spatial regions? Outside of the laboratory, visual scenes typically contain multiple objects. As a result, the visual system must integrate motion measurements originating from the same object while segmenting motion measurements arising from different objects. The boundaries of objects are frequently indicated by shape discontinuities such as corners and terminators (Shimojo, Silverman,

& Nakayama, 1989). The presence of such discontinuities might therefore signal when motion measurements should be segregated across edges. To test this hypothesis, the revolving diamond stimulus described previously was modified in several ways (Lorenceau & Shiffrar, 1992; Shiffrar & Lorenceau, 1996). First, the luminance contrast of the endpoints, or terminators, of each visible line segment was varied while the luminance of the central portion of each line segment was held constant. When terminator visibility was low, because terminators were presented at low luminance contrast relative to the background, then observers interpreted the four translating edges as part of one coherently revolving diamond. However, as terminator contrast increased and terminators became more visible, the four edges of the diamond appeared to move independently and the veridical perception of a coherent diamond was lost (Lorenceau & Shiffrar, 1992). In another manipulation of terminator visibility, the translating diamond stimulus was rendered near-isoluminant with its background for each observer. Under this condition, the translating diamond appeared to be coherent. As the luminance contrast of the diamond increased, it increasingly appeared to move incoherently (Shiffrar & Lorenceau, 1996). These observations, along with experimental data (Lorenceau & Shiffrar, 1992; Shiffrar & Lorenceau, 1996), indicate that the visibility of the ends of line segments influences whether motion information is integrated or segmented across those segments—that is, whether the visual system interprets line segments as pieces of coherent objects or as independent fragments.

In another test of the discontinuity hypothesis described earlier, the visibility of the terminators in the revolving diamond stimuli was manipulated with positional noise. In the original display (Fig. IV.76-4b) the translating diamond was seen through four rectangular apertures. In this noise manipulation, a variable amplitude sine wave was added to each aperture edge (Fig. IV.76-4c). When the diamond revolved behind these apertures, the visible terminators of each edge followed irregular paths. This addition of the positional jitter to the terminators increased the perception of coherent diamond motion (Lorenceau & Shiffrar, 1992). In other words, decreased terminator visibility, per se, facilitated motion integration across space. Another manipulation of terminator visibility took advantage of the fact that positional uncertainty increases in peripheral vision (Burbeck & Yap, 1990). When observers view the original diamond stimulus translating behind apertures (Fig. IV.76-4b) with their central vision, the edges of the diamond appear to move independently. However, when observers view this same stimulus in their visual periphery, the diamond appears to move coherently, suggesting that motion integration across space has taken place. In sum, the balance between motion integration and motion segmentation depends on those factors that determine terminator visibility.

WHY IS THE APERTURE PROBLEM INTERESTING AND RELEVANT TO VISION RESEARCH?

The visual interpretation of moving objects requires the visual system to perform a balancing act of sorts. On one hand, the visual system must group motion measurements together in order to determine an object's motion. Grouping is necessary because initial motion measurements can record only the velocities of small regions of an object in a manner that is often ambiguous as a result of the aperture problem. Yet segmentation is also needed to confine grouping within objects. Without segmentation, all of the motion signals in a scene would be grouped together into one giant, and uninformative, motion soup.

One of the take-home messages from the illusions and studies described here is that the visual system solves this balancing act between segmentation and integration by ignoring the theoretical separation of form and motion processes that has been proposed in numerous classic models of the visual system (e.g., Livingstone & Hubel, 1988). In other words, the aperture problem is a problem with the extent to which the visual system is strictly modular and motion analyses are conducted independently of form analyses. If motion analyses cannot take advantage of disambiguating form cues, then all motion signals must be weighted equally and the aperture problem because a very big problem indeed. However, the studies summarized here indicate that the visual system can be understood as sidestepping many aspects of the aperture problem by simply considering motion and form cues together. In this way, simple shape discontinuities can determine how motion measurements are interpreted. It follows that, while theoretical models of the visual system may be modular, the visual analyses themselves may not be.

Traditionally, the visual sciences have avoided the study of dynamic image discontinuities or features, such as endpoints and corners, because such discontinuities are inherently nonlinear. As such, they are often costly to compute under naturally noisy conditions. For example, it can be difficult to solve the correspondence problem for a dynamic two-dimensional feature presented within noise or at low-luminance contrast because the number of potential false correspondences increases with increasing noise and decreasing contrast. However, the studies and illusions described here suggest that simple shape discontinuities determine whether the visual system selects a holistic, object-based interpretation or a segmented, fragment-based interpretation of a visual scene.

VARIANTS AND RELEVANT PARAMETERS FOR THE EFFECT

By what mechanism might simple shape discontinuities determine when local velocity estimates are combined? In a study examining this question, observers reported the perceived direction of translating oblique line segments across variations of line length, luminance contrast, and display duration (Lorenceau, Shiffrar, Wells, & Castet, 1993). Systematic errors in the perceived direction of line translation were consistent with a simple model of motion perception involving two distinct classes of receptors that have been previously defined by electrophysiological studies (e.g., Hubel & Wiesel, 1968). One class of receptors measures the velocity of straight edges quickly while the second class measures the velocity of discontinuities more slowly. These two types of receptors appear to compete within a simple averaging process, weighted according

to the perceptual saliency of the two feature types and the temporal duration of displays (Castet, Lorenceau, Shiffrar, & Bonnet, 1993). Subsequent neurophysiological evidence supports the proposal that the motion processing system takes into account the degree of uncertainty associated with motion measurements from different form cues (Pack et al., 2004; Pack, Livingstone, Duffy, & Born, 2003). Less ambiguous velocity estimates dominate the interpretation of more ambiguous velocity estimates, especially as stimulus duration increases, even when those velocity estimates are separated spatially (Shiffrar, Li, & Lorenceau, 1995).

While discontinuities may control motion integration, an unrestricted reliance on them would result in image segmentation errors. This might occur, for example, if a discontinuity from one object dictated how the motion of a different object were interpreted. One way in which the visual system appears to limit this problem is by facilitating the integration of motion signals across collinear edges (Barchilon Ben-Av & Shiffrar, 1995). Such structural biases in motion integration may assist the visual system in successfully juggling the need for integration within but segmentation across objects.

ADDITIONAL RESEARCH AND THEORIES

Does the visual system balance the motion integration and segmentation processes across all categories of stimuli in the same way? Some researchers have argued that the analysis of human movement differs significantly from the visual analysis of nonhuman, or object, movement (e.g., see Shiffrar, 2011, for review). This raises the question of whether the integration of human motion and nonhuman motion across apertures differ. To examine this question, dynamic line drawings of two moving objects (translating car, opening and closing scissors) and a walking person were viewed through a set of stationary apertures. The apertures were constructed so that only straight edges (i.e., no corners or endpoints) were visible. Naïve observers simply viewed one of these three stimuli and reported what they saw (Shiffrar, Lichtey, & Heptulla Chatterjee, 1997).

When observers viewed moving, nonhuman objects (the scissors or cars) through apertures, they were unable to recognize these objects. Indeed, observers reported perceiving sets of line segments moving incoherently when they viewed these objects. Such descriptions suggest that observers could not integrate the cues to object motion across the apertures. This finding is consistent with the rotating and translating diamond stimuli described previously (Fig. IV.76-4). However, when observers viewed the walking human figure through apertures, they readily and accurately identified the walker. Recognition of the walker was orientation specific and robust across a range of stimulus durations (Shiffrar et al., 1996). Yet recognition of human gait was possible only when the spatial and temporal parameters of the display were consistent with biomechanically plausible walking speeds. These results suggest that there may be significant differences in how the visual system integrates human and nonhuman motion across apertures.

REFERENCES

Adelson, E. H., & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, *300*, 523–525.

Barchilon Ben-Av, M., & Shiffrar, M. (1995). Disambiguating velocity estimates across image space. *Vision Research*, *35*, 2889–2895.

Burbeck, C. A., & Yap, Y. L. (1990). Two mechanisms for localization? Evidence for separation dependent and separation independent processing of position information. *Vision Research*, *30*, 739–750.

Castet, E., Lorenceau, J., Shiffrar, M., & Bonnet, C. (1993). Perceived speed of moving lines depends on orientation, length, speed and luminance. *Vision Research*, *33*, 1921–1936.

Hildreth, E. (1984). *The measurement of visual motion*. Cambridge, MA: MIT Press.

Hubel, D., & Wiesel, T. (1959). Receptive fields of single neurons in the cat's striate cortex. *Journal of Physiology*, *148*, 574–591.

Hubel, D., & Wiesel, T. (1968). Receptive fields and functional architecture of the monkey striate cortex. *Journal of Physiology*, *195*, 215–243.

Hubel, D. H. (1982). Exploration of the primary visual cortex, *Nature*, *299*, 515–524.

Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science*, *240*, 740–749.

Lorenceau, J., & Shiffrar, M. (1992). The influence of terminators on motion integration across space. *Vision Research*, *32*, 263–273.

Lorenceau, J., Shiffrar, M., Wells, N., & Castet, E. (1993). Different motion sensitive units are involved in recovering the direction of moving lines. *Vision Research*, *33*, 1207–1217.

MacNulty, A. S. (1945). The Renaissance and its influence on English medicine, surgery, and public health. *The British Medical Journal*, *2*, 755–759.

Majaj, N. J., Carandini, M., & Movshon, J. A. (2007). Motion integration by neurons in macaque MT is local, not global. *The Journal of Neuroscience*, *27*, 366–370.

Movshon, J. A., Adelson, E. H., Gizzi, M. S., & Newsome, W. T. (1985). The analysis of moving visual patterns. In C. Chagas, R. Gattas, & C. G. Gross (Eds.), *Pattern recognition mechanisms* (pp. 117–151). Rome: Vatican Press.

Movshon, J.A., Thompson, I. D., & Tolhurst, D. J. (1978). Receptive field organization of complex cells in the cat's striate cortex. *Journal of Physiology*, *283*, 79–99.

Pack, C. C., Gartland, A. J., & Born, R. T. (2004). Integration of contour and terminator signals in visual area MT of alert Macaque. *The Journal of Neuroscience*, *24*, 3268–3280.

Pack, C. C., Livingstone, M. S., Duffy, K. R., & Born, R. T. (2003). End-stopping and the aperture problem: Two-dimensional motion signals in the Macaque V1. *Neuron*, *39*, 671–680.

Saxe, J. G. (1855). *Poems by John Godfrey Saxe*. Boston: Ticknor and Fields.

Shiffrar, M. (2011). People watching: Visual, motor, and social processes in the perception of human movement. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*, 68–78.

Shiffrar, M., Li, X., & Lorenceau, J. (1995). Motion integration across differing image features. *Vision Research*, *35*, 2137–2146.

Shiffrar, M., Lichtey, L., & Heptulla Chatterjee, S. (1997). The perception of biological motion across apertures. *Perception & Psychophysics*, *59*, 51–59.

Shiffrar, M., & Lorenceau, J. (1996). Increased motion linking across edges with decreased luminance contrast, edge width and duration. *Vision Research*, *36*, 2061–2067.

Shiffrar, M., & Pavel, M. (1991). Percepts of rigid motion within and across apertures. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 749–761.

Shimojo, S., Silverman, G. H., & Nakayama, K. (1989). Occlusion and the solution to the aperture problem for motion. *Vision Research*, *29*, 619–626.

Stumpf, P. (1911) Über die Abhangigkeit der visuellen Bewegungsrichtung und negativen Nachbildes von den Reizvorgangen auf der Netzhaut. *Zeitschrift fur Psychologie*, *59*, 321–330.

Todorović, D. (1996) A gem from the past: Pleikart Stumpf's (1911) anticipation of the aperture problem, Reichardt detectors, and perceived motion loss at equiluminance. *Perception*, *25*, 1235–1242.

Ullman, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.

Wallach, H. (1935). Uber visuell wahrgenommene Bewegungsrichtung. *Psychologische Forschung*, *20*, 325–380.

Wuerger, S., Shapley, R., & Rubin, N. (1996). "On the visually perceived direction of motion" by Hans Wallach: 60 years later. *Perception*, *25*, 1317–1367.

Chapter 77
Paths of Apparent Human Motion Follow Motor Constraints

Maggie Shiffrar and Christina Joseph

PRESENTATION AND DESCRIPTION OF THE EFFECT

What Is Apparent Motion?

Apparent motion is typically understood as an optical illusion because it defines the perception of motion where none exists. Apparent motion is widely studied because it represents an excellent example of the constructive nature of motion perception and, not unrelatedly, played a central role in the foundation of the perceptual sciences. The scientific significance of apparent motion is typically traced back to a 1910 train trip by Max Wertheimer. According to legend, during some point on this trip, Wertheimer played with a child's toy, known as a flip book, that he had brought as a gift for an upcoming visit (Steinman, Pizlo, & Pizlo, 2000). As Wertheimer flipped through the static pages of that book, he deeply considered the meaning of the resulting illusory percepts of motion. Wertheimer later elaborated upon his insights from that flip book in his classic 1912 article; an article frequently cited as the founding moment of Gestalt psychology (Ash, 1995). Besides describing his observations from the flip book, and providing a theoretical interpretation of them, Wertheimer drew the revolutionary conclusion that percepts are qualitatively different from low-level stimulus features. In other words, apparent motion, or the illusory perception of movement from stationary images, indicates that perception does not simply result from a summation of information from the stimulus array—it must be something more.

Anyone who has played with an old-fashioned flip book has probably noticed that sequentially presented static pictures can give rise to the perception of movement. Whether we perceive apparent motion depends on time and space—namely, how rapidly static images are presented and the distance between those images. Consider, for example, a very simple case of two stationary dots flashed in alternation (see Fig. IV.77-1a–b). If the frames depicting the static dots are flashed rapidly (roughly 50 msec or less, depending on the spatial separation), then observers tend to perceive the stimulus accurately as two briefly presented stationary dots. However, if these static images are presented more slowly (roughly 50 to 250 msec), then observers tend to report the perception of smooth apparent motion, that is, the perception of a single dot translating (Braddick, 1980). If such a series of static images were presented even more slowly, then the perception of apparent motion would eventually be lost and observers would again perceive static dots. These temporal windows should not be taken as fixed values because apparent motion perception depends on several factors, including display complexity (DeSilva, 1926; Mather, 1988).

How Did It Get There?

One of the many fascinating features of apparent motion is the construction of perceived paths of motion (e.g., Kim, Feldman, & Singh, 2013; Shepard & Lare, 1983; Yantis & Nakama, 1998). If a stimulus depicts only the starting location and the ending location of an object, then how does our visual system determine the path that connects those two locations? Obviously there can exist an infinite number of paths that can connect any two points (Fig. IV.77-1c). So how does our visual system compute a single path from an infinite number of possible paths? Classic research indicates that our visual systems solve this problem by defaulting to the shortest possible path of apparent motion (Burt & Sperling, 1981).

An interesting example of this shortest path constraint can be seen in the wagon wheels sometimes shown in old Western movies. In and of themselves, movies constitute a wonderful example of apparent motion because while films consist of an extended series of rapidly displayed static images, we readily perceive movement in films. Indeed, it can be a challenge to explain to students that the motion we perceive in the cinema and on monitors is not physically present but is rather a product of our visual system's reconstruction of the outside world. The wagon-wheel illusion can be seen in films and on TV (Marr, 1982; Purves, Paydarfar, & Andrews, 1996) when the spokes of a wagon wheel rotate very quickly, such that across presentation frames the spokes rotate more than the distance separating those spokes, and, as a result, the shortest path of motion between pairs of spokes is backward rather than forward (Fig. IV.77-1d). When this occurs, the wagon wheels appear to rotate backward while the wagon to which they are attached moves quickly forward. This illusion illustrates that the visual system relies heavily on the shortest path

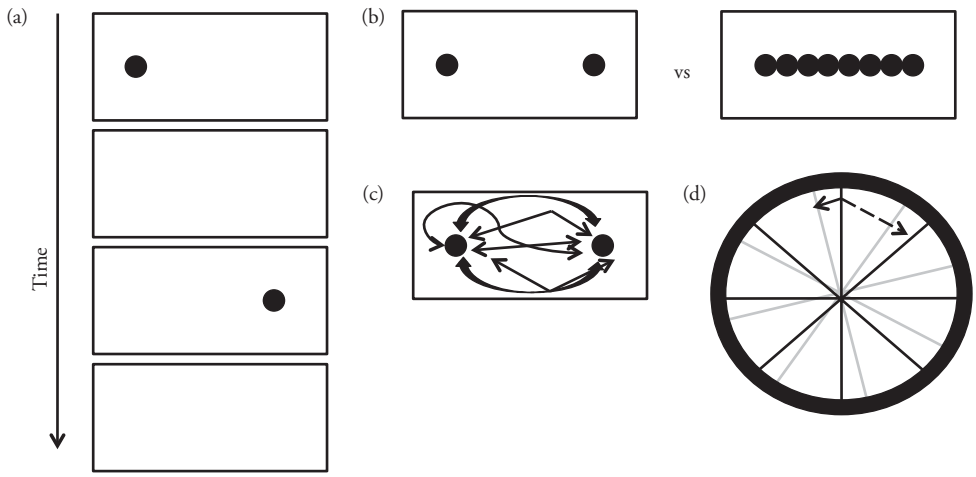


Figure IV.77-1. (a) An illustration of an apparent motion display. In this case, four static images are repeatedly briefly presented in sequential order. On the first image, a dot is located on the left and on the third image a dot is located on the right. (b) Two possible perceptual interpretations of stimulus A. Observers tend to report the perception of two stationary dots that flash in alternation (left image) or the perception of one moving dot that translates smoothly back and forth (right image). Time plays a large role in determining when such apparent motion is perceived. (c) A few examples of the infinite number of paths that two dots in apparent motion could, theoretically, traverse. The visual system typically selects the shortest possible path of apparent motion. (d) The wagon-wheel illusion. Two overlapping views of a wagon wheel rotating clockwise so that the spokes appear in the location of the black spokes first and the gray spokes second. However, because of the shortest path constraint, observers erroneously perceive the wheel as rotating counterclockwise. (Figures by Maggie Shiffrar.)

constraint when calculating the paths of objects in apparent motion.

Apparent Motion of the Human Body

If visual percepts of apparent motion were computed in the same way for every possible type of displaced stimulus, then the shortest path constraint should consistently dictate the perception of apparent motion. This prediction was beautifully described by one of the most important researchers in the study of apparent motion, Roger N. Shepard, when he stated in 1984 that

> There evidentially is little or no effect of the particular object presented. The motion we involuntarily experience when a picture of an object is presented first in one place and then in another, whether the picture is of a leaf or of a cat, is neither a fluttering drift nor a pounce, it is, in both cases, the same simplest rigid displacement. (p. 426)

Subsequent research, however, has identified an important exception to this rule: the visual perception of apparent motion of the human body. By definition, the shortest path that connects any two points is a straight line. However, the paths of motion that our bodily limbs traverse frequently violate this shortest path constraint. For example, bodily limbs follow curvilinear rather than rectilinear motion trajectories because human motion is constrained by the joints about which our limbs move. When people act, their movements are also constrained by the solidity of their bodies because one solid limb cannot pass through another. As a result, limbs must move around bodies, rather than through them, which again necessitates paths of motion that are longer than the shortest possible path. This raises the question of whether percepts of the human body in apparent motion displays follow the shortest possible path of motion or the longer, biomechanically feasible paths of motion.

In the first test of this question, naïve observers viewed photographs of a human model in different postures. These postures were selected so that the biomechanically possible paths of limb motion conflicted with the shortest possible paths (Shiffrar & Freyd, 1990). An illustration of a stimulus, shown in Figure IV.77-2a, consists of two images of a person. In the first, the standing woman points her right arm up and back, while in the second she points that same arm down and back. In between these two images are blanks of variable durations. The shortest path connecting the two arm positions (Fig. IV.77-2b) would require the arm to move behind the woman's body (a biomechanically impossible action given the rotational limitations of the shoulder). Conversely, a biomechanically plausible trajectory of apparent motion would be longer and involve the arm rotating naturally forward toward the front of the body.

When observers view such stimuli, their perceived paths of motion change with the amount time between the onset of one image and the onset of the next image (known as the stimulus onset asynchrony (SOA). When such image pairs appear in rapid alternation, observers report seeing the shortest, physically impossible motion path, as predicted by the shortest path constraint. However, as the rate of picture presentation slows, observers are increasingly likely to see longer paths of apparent motion that are consistent with the ways in which the human body moves (Shiffrar & Freyd, 1990). In an important control condition, these same observers also viewed images of inanimate control objects, such as scissors, erasers, cars, and clocks, positioned so that they subtended the same angles and were separated by the same distances (in degrees of visual angle) as the human body postures. When these inanimate

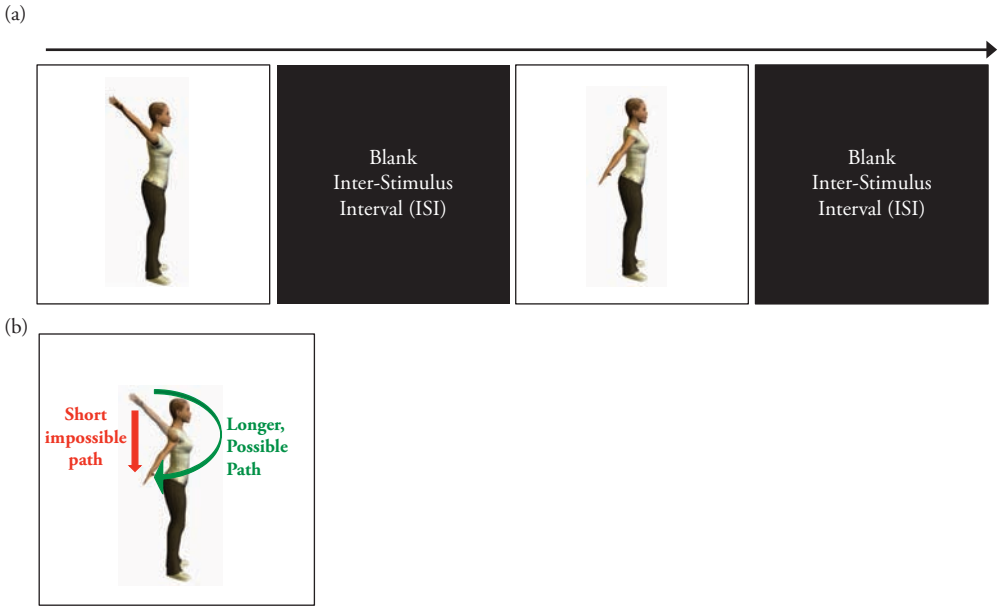


Figure IV.77-2. (a) An apparent motion stimulus depicting a woman with two different arm positions. (b) The shortest path connecting the two arm positions is a physically impossible straight path behind the woman's back. The typical, physically possible path is curved in front of the woman's body. When these two images are flashed in rapid alternation, the shortest, physically impossible path is perceived. When the presentation rate is slowed, the physically possible path of motion is observed. When the shortest possible path of motion is biomechanically plausible, then observers report the perception of short paths of apparent motion at all temporal display rates. (Figures by Maggie Shiffrar.)

objects were presented at the same rates of alternation as the human body stimuli, observers consistently reported seeing the shortest paths of apparent motion at all displays rates, even though these short paths were physically impossible. Thus inanimate objects in apparent motion tend to traverse paths that are consistent with the shortest path constraint.

Importantly, when viewing apparent motion displays of a human model positioned so that the shortest movement path is also a biomechanically plausible path, observers always report seeing that short and biomechanically plausible path of motion (Shiffrar & Freyd, 1993). Thus subjects do not simply report the perception of longer paths with slower presentation rates. Instead, in the case of the human body, observers report increasingly accurate paths of apparent motion with slower display rates.

The results of follow-up studies suggest that the perception of paths of apparent motion that are consistent with the biomechanical limitations of human movement extends beyond limb movements about an actor's own body and also includes human limb movements about inanimate objects (Heptulla Chatterjee, Freyd, & Shiffrar, 1996). For example, at slow rates of alternation, static images of an actor with her arm positioned on either side of a brick will give rise to the perception of apparent arm motion around the brick rather than through it. Such findings suggest that the influence of biomechanical constraints on the visual perception of human movement is general and robust in that it can be found under multiple conditions. Interestingly, the global structure of the human body is necessary for the perception of biomechanically plausible paths of apparent motion at slow display rates (Heptulla-Chatterjee et al., 1996). Thus, for example, when only two limbs are depicted in apparent motion (and the rest of the body is erased), observers consistently report, at all display rates, the perception of one limb moving through another, that is, following the shortest path. Conversely, when a posed wooden mannequin replaces more realistic depictions of a human body, observers report the perception of biomechanically plausible paths of apparent human motion at slow display rates even though the underlying stimulus is wooden. These results indicate that the global hierarchy of position and orientation cues consistent with the human form provides sufficient form information for the perception of biomechanically plausible paths of apparent human motion. This conclusion is consistent with the well-known effect that human motion can be readily perceived even when the movements of an actor have been reduced to the movements of a few point-lights strategically positioned on the actor's head and major joints (Johansson, 1973).

WHY THIS EFFECT IS INTERESTING AND RELEVANT TO VISION RESEARCH

Studies of the visual perception of apparent human motion advance our understanding of vision in some key ways. The first relates to the role of prediction. The speed of neural processing is slow relative to the speed of many changes in the outside environment. As a result, for example, whenever someone tosses a friend a ball, by the time the receiver's visual system has determined the ball's location, the ball is no longer there. So how do we successfully interact with a rapidly changing world when our visual system is relatively slow? One explanation is that the neural processes underlying visual perception are fundamentally predictive in that they compute not the current status of a stimulus

but rather where that stimulus will be and how events will unfold in the near future. Exactly how and where such compensatory prediction takes place has been an issue of much debate and research (e.g., Nijhawan, 2009; Perrett, Xiao, Barraclough, Keysers, & Oram, 2009). However perceptual prediction takes place, it is nonetheless reasonable to assume that better predictive ability should lead to more accurate percepts.

By almost any measure, the human body represents a special kind of visual stimulus. As inherently social beings, people have extensive visual experience with and pay an extraordinary amount of attention to the actions of other people. In this way, typical humans can be understood as experts in the perception of human motion. This expertise should facilitate predictive processes. Prediction and apparent motion can be understood as related processes because the "filling-in" or construction of perceived paths of apparent motion requires a prediction of where and how a stimulus might move. To the extent that human observers are visual experts in human motion, it is important to remember that such perceptual expertise is constrained. While people have extensive experience watching other people perform physically possible movements, they have very little experience watching physically impossible movements. Because people have more experience watching possible human movements than impossible human movements, observers should be better able to predict or fill-in possible paths of human movement in apparent motion displays.

This idea was tested in a priming study in which observers viewed pairs of human body postures that were linked by apparent motion or presented so as to appear static. Immediately after each stimulus pair, observers then viewed another pair of static body postures and reported whether these two static postures were the same or different from one another. The key manipulation was whether these judged body postures fell within (i.e., were consistent with) the path of apparent human motion seen immediately beforehand. Observers were significantly faster—that is, they showed priming—when judging body postures that would have fallen within the path of apparent human motion just seen (Kourtzi & Shiffrar, 1999). Body shape similarity judgments were relatively slow for body postures that would not have been seen in the path of apparent motion just viewed (Kourtzi & Shiffrar, 1999). These results suggest that observers automatically generate visual representations of body postures not actually seen but only predicted or filled-in during the perception of apparent human motion. Taken together, these results suggest that perceptual prediction based on visual experience plays a key role in visual analyses of human movement.

The finding that the visual perception of apparent human motion is shaped by biomechanics raises another question; namely, how does this shaping occur? Does the visual system rely on the observer's conscious conceptual knowledge of biomechanics? Or are visual processes shaped by the observer's own motor experience? The previous discussion focused on the role that visual expertise might play in shaping our perception of apparent motion. But motor expertise provides another source of information that our visual systems might use to facilitate analyses of human action.

Unlike any other category of visual motion, human movement is the only stimulus that human observers can accurately reproduce with their own bodies. We can move like other people move, but we cannot move in exactly the same ways as drifting clouds, wind-blown trees, or crashing waves. If our visual system could take advantage of the information available in our motor system during visual analyses of human action, then this might enhance visual sensitivity to other people's actions. The hypothesis that visual processes and motor processes are interactive, rather than independently modular, is known as perception-action coupling or embodied perception (e.g., Prinz, 1997; Shiffrar & Pinto, 2002; Wilson, 2001, 2002). Some researchers have argued that the visual system is so inherently dependent on the motor system that the only condition under which observers perceive movement accurately is when that movement is something that the observer's own motor system could have generated (e.g., Viviani, 2002). The theory of perception-action coupling provides an intriguing explanation of why the visual perception of humans and objects in apparent motion might differ; namely, that motor processes might contribute to the visual analysis of the human body in a manner that differentiates percepts of apparent human motion from percepts of apparent object motion.

A first step in testing this potential explanation is to determine whether neural activity in the motor systems of stationary observers is elevated during the perception of apparent human motion. To answer this question, a brain imaging study was conducted in which passive, unmoving observers viewed apparent motion displays of human motion and object motion across a spectrum of temporal display rates (Stevens, Fonlupt, Shiffrar, & Decety, 2000). Changes in brain activity, as measured by positron emission tomography, were assessed across differences in image type (object vs. human) and display rate. As in previous human apparent motion studies, the picture pairs in the apparent motion displays used here were constructed so that physically plausible paths of motion were longer than the shortest, physically impossible paths of motion (as illustrated in Fig. IV.77-2). Recall that for depictions of the human body in apparent motion displays, observers typically report the perception of biomechanically plausible paths of motion at slow display rates and the perception of impossible paths at rapid display rates (Shiffrar & Freyd, 1990). This perceptual effect allowed for the differentiation of the neural areas active during the perception of possible and impossible paths of human movement while keeping the bodily shape and posture cues constant across the two type of perceived motion. The results of this study indicated increased levels of brain activity in the motor cortex of passive observers during the perception of biomechanically possible paths of apparent human motion. No change in motor system activity, relative to rest, was found during the perception of impossible paths of apparent human motion or during the perception of apparent object motion at either rapid or slow presentation rates (Stevens et al., 2000). Such results indicate that motor system activity is associated with the perception of paths of apparent motion that are consistent with the ways in which people move.

VARIANTS AND RELEVANT PARAMETERS FOR THE EFFECT

Two basic parameters define the temporal display rates used in studies of apparent motion. The stimulus duration is the duration during which each image is shown. The interstimulus interval is the amount of time between each image. These two parameters can be combined (stimulus duration + interstimulus interval) to create the SOA, or the time from the onset of one image until the onset of the next image. The apparent human motion studies described here all varied the SOA because previous studies had demonstrated a clear dependence of SOA on perceived paths of apparent motion (e.g., Shepard, 1984). Across studies (Heptulla-Chatterjee et al., 1996; Shiffrar & Freyd, 1993) there does not appear to be a fixed or special SOA at which percepts of apparent motion of the human body shift from impossible to possible paths of human movement. Instead, informal observation suggests that human actions that require more time to produce (e.g., lifting one foot about the other) require longer SOAs for the perception of biomechanically plausible paths than actions that can be rapidly produced (e.g., moving one finger about another finger). In other words, there appears to be a link between the amount of time needed to physically perform a particular human action and the SOA at which observers predominantly perceive the physically possible motion trajectory for that action. Thus it does not appear to be the case that the processing time (or SOA) needed for the perception of possible paths of human movement reflects the time needed to access some conceptual knowledge store. Rather it seems that the temporal constraints on the perception of any biomechanically veridical path of apparent motion might reflect the time that the observer would need to physically produce that same motor trajectory.

Research examining the impact of motor constraints on visual perception is consistent with the temporal linkage hypothesized here. Fitts's law, for example, is one of the best-documented constraints on motor production. It states that the time needed to move as quickly as possible between two targets is a function of the distance between the targets and target width (Fitts, 1954). Research on the visual perception of apparent finger motion between pairs of targets of different widths and separation distances indicates that Fitts's law also applies to the perception of action in apparent motion displays (Grosjean, Shiffrar, & Knoblich, 2007). In other words, the visual perception of human movement in apparent motion displays is consistent with the temporal characteristics of motor production. This further supports the hypothesis that motor processes shape the visual perception of apparent motion so that perceived paths of human motion are consistent with the biomechanical constraints on movement production.

ADDITIONAL RESEARCH AND THEORIES

If the perception of human action in apparent motion displays depends on the observer's own motor experience, then one would expect to find that idiosyncratic differences in motor ability should result in idiosyncratic differences in the visual perception of human movement. In a test of this prediction, two individuals born without hands viewed apparent motion displays depicting hand rotations (Fig. IV.77-3) at different SOAs (Funk, Shiffrar, & Brugger, 2005). While these two individuals were well matched on many different measures, including intelligence and physical appearance, they differed in whether or not they experienced phantom hands. Control observers, all born with hands, saw the same apparent motion hand displays at the same temporal rates. The question at the heart of this study was whether observers born without hands would perceive rotating hands in apparent motion displays as objects (always the shortest path of apparent motion) or as bodies (SOA-dependent paths of apparent motion).

As before, all observers viewed pairs of static images of sequentially presented hands under conditions of apparent motion and at a variety of display rates. Again, as before, the shortest possible path of apparent hand rotation was physically impossible. The results demonstrated that control observers all increasingly reported the perception of biomechanically possible paths of apparent hand rotation with increases in SOA (i.e., slower presentation rates). The individual born without hands but who experienced phantom hands also reported the same temporally dependent pattern of apparent motion percepts: increases in SOA led to increased percepts of biomechanically plausible paths of apparent hand rotation. However, the individual born without hands and lacking the experience of phantom hands consistently reported the perception of physically impossible paths of apparent hand rotation at all SOAs (Funk et al., 2005). These intriguing data suggest that the experience of phantom limbs, in the absence of physical limbs, is sufficient to shape the visual perception of apparent limb motion. This likely occurs through internal representations of the observer's own body, known as body schema and located in the parietal lobe (Funk et al., 2005), which are known to contribute to visual comparisons between the self and observed others during the visual analysis of other people's actions (Reed, 2002). In short, it appears that visual perception of biomechanically plausible paths of apparent human motion depends on motor experience independent of whether that experience comes from physical limbs or phantom limbs.

In conclusion, the illusion of apparent motion is shaped not only by factors such as space and time but also by the observer's motor experience with his or her own body. As

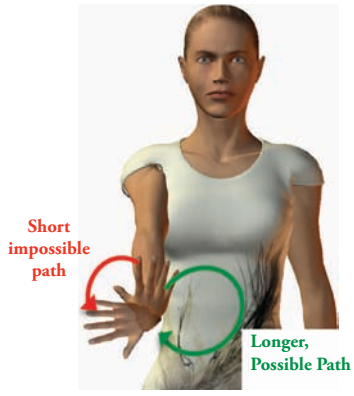


Figure IV.77-3. Illustration of two possible paths of apparent hand rotation. (Adapted from Funk et al., 2005.)

reviewed, observers perceive biomechanically possible path of human motion when apparent motion displays are presented at temporal rates that are consistent with the amount of time needed to physically perform the particular action being observed. When an person's action is presented in apparent motion displays more quickly than the person could physically perform that action, then the human visual system appears to default back to the shortest path constraint, and, as a result, observers perceive short and physically impossible paths of apparent motion. Visual experience may also contribute to the perception of biomechanically plausible paths of apparent human motion by facilitating predictive processes. Taken together, studies of the robust illusion of apparent motion shed light on how our visual system interprets and represents the physical and social worlds.

REFERENCES

Ash, M. G. (1995). *Gestalt psychology in German culture, 1890–1967: Holism and the quest for objectivity*. Cambridge, UK: Cambridge University Press.

Braddick, O. J. (1980). Low-level and high-level process in apparent motion. *Philosophical Transactions of the Royal Society of London*, *290*, 131–151.

Burt, P., & Sperling, G. (1981). Time, distance and feature trade-offs in visual apparent motion. *Psychological Review*, *7*, 171–195.

DeSilva, H. R. (1926). An experimental investigation of the determinants of apparent visual motion. *Journal of Experimental Psychology*, *37*, 469–501.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, *47*, 381–391.

Funk, M., Brugger, P., & Shiffrar, M. (2005). Hand movement observation by individuals born without hands: Phantom limb experience constrains visual limb perception. *Experimental Brain Research*, *164*, 341–346.

Heptulla Chatterjee, S., Freyd, J., & Shiffrar, M. (1996). Configural processing in the perception of apparent biological motion. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 916–929.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, *14*, 201–211.

Kim, S.-H., Feldman, J., & Singh, M. (2013). Perceived causality can alter the perceived trajectory of apparent motion. *Psychological Science*, *24*, 575–582.

Kourtzi, Z., & Shiffrar, M. (1999). Dynamic representations of human body movement. *Perception*, *28*, 49–62.

Marr, D. (1982). *Vision*. San Francisco, CA: Freeman.

Mather, G. (1988). Temporal properties of apparent motion in subjective figures. *Perception*, *17*, 729–736.

Nijhawan, R. (2009). Visual prediction: Psychophysics and neurophysiology of compensation for time delays. *Behavioral and Brain Sciences*, *31*, 179–239.

Perrett, D. I., Xiao, D., Barraclough, N. E., Keysers, C., & Oram, M. W. (2009). Seeing the future: Natural image sequences produce "anticipatory" neuronal activity and bias perceptual report. *Quarterly Journal of Experimental Psychology*, *62*, 2081–2104.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, *9*, 129–154.

Purves, D., Paydarfar, J. A., & Andrews, T. J. (1996). The wagon wheel illusion in movies and reality. *Proceedings of the National Academy of Sciences USA*, *93*, 3693–3697.

Reed, C. L. (2002). What is a body schema? In W. Prinz & A. Meltzoff (Eds.), *The imitative mind: Development, evolution, and brain bases* (pp. 233–243). Cambridge, UK: Cambridge University Press.

Shepard, R. N. (1984). Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review*, *91*, 417–447.

Shepard, R. N., & Zare, S. L. (1983). Path guided apparent motion. *Science*, *220*, 632–634.

Shiffrar, M., & Freyd, J. J. (1990). Apparent motion of the human body. *Psychological Science*, *1*, 257–264.

Shiffrar, M., & Freyd, J. J. (1993). Timing and apparent motion path choice with human body photographs. *Psychological Science*, *4*, 379–384.

Shiffrar, M., & Pinto, J. (2002). The visual analysis of bodily motion. In W. Prinz & B. Hommel. *Common mechanisms in perception and action* (pp. 381–399). Attention and Performance 19. Oxford: Oxford University Press.

Steinman, R. M., Pizlo, Z., & Pizlo, F. (2000). Phi is not beta, and why Wertheimer's discovery launched the Gestalt revolution. *Vision Research*, *40*, 2257–2264.

Stevens, J. A., Fonlupt, P., Shiffrar, M., & Decety, J. (2000). New aspects of motion perception: Selective neural encoding of apparent human movements, *NeuroReport*, *11*, 109–115.

Viviani, P. (2002). Motor competence in the perception of dynamic events: A tutorial. In W. Prinz & B. Hommel (Eds.), *Common mechanisms in perception and action* (pp. 406–442). Attention and Performance 19. Oxford: Oxford University Press.

Wilson, M. (2001). Perceiving imitatable stimuli: Consequences of isomorphism between input and output. *Psychological Bulletin*, *127*, 543–553.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review*, *9*, 625–636.

Yantis, S., & Nakama, T. (1998). Visual interactions in the path of apparent motion. *Nature Neuroscience*, *1*, 508–512.

Chapter 78
The Motion Standstill Illusion

George Sperling, Son-Hee Lyu, Chia-Huei Tseng, and Zhong-Lin Lu

INTRODUCTION

In the motion standstill illusion, a pattern that is moving quite rapidly is perceived as being absolutely motionless, and yet its details are not blurred but clearly visible. This is a very unusual phenomenon, and it takes very special stimuli to demonstrate it. To understand why motion standstill occurs, it is useful to recognize that even when we keep as motionless as possible, small, involuntary movements of the body and eyes cause the retinal image to make random movements across the retina that are large in relation to details that can easily be resolved (Rossi & Roorda, 2010). During active vision in normal situations, as in walking, the image on the retina continuously undergoes involuntary rapid random excursions that are enormously larger than details of shape, texture, color, and depth that can be easily resolved in the randomly moving retinal image that, despite its uncontrolled random movement, is perceived as being stable (Steinman, Cushman, & Martins, 1982).

Theoretically, motion standstill is a very important phenomenon. It demonstrates that the job of the shape, texture, color, and depth systems is to extract a stable, static percept from the normal, randomly moving visual input. When we do perceive motion, it is the result of a computation by the motion system. When devious experimental manipulations silence the motion system under circumstances that enable the other systems to still operate, we perceive only the stable static image that the shape, texture, color, and depth systems derive from the moving retinal image.

FIVE WAYS TO PRODUCE THE MOTION STANDSTILL ILLUSION

1. Motion standstill can occur with rapidly moving fine-detail images that have been constructed so that the spatiotemporal resolution of the shape-texture-object perception systems exceeds that of the first-order motion system and, incidentally, all other motion systems as well (Sperling, Lyu, & Kim, 2002).
2. Motion standstill occurs with moving color stimuli that are constructed to be equiluminant to silence the first-order motion system, that have insufficient texture to activate the second-order motion system, and that are isosalient to silence the third-order motion system. The requirements of equiluminance and isosalience place severe constraints on the range of possible images but nevertheless leave enough freedom to enable representation of objects that are completely invisible to the motion systems (and therefore cannot produce the sensation of motion) but are perfectly visible to shape, texture, color, and depth systems (Lu, Lesmes, & Sperling, 1999b).
3. Motion standstill occurs when the stereodepth resolution of the depth perception system exceeds that of the third-order motion system (Tseng, Gobell, Lu, & Sperling, 2006).
4. Motion of stimuli that stimulate primarily blue cones selectively disadvantage the motion systems (Cavanagh, Tyler, & Favreau, 1984).
5. Various motion adaptation procedures (prolonged viewing of high-contrast moving stimuli) adapt/fatigue the motion systems sufficiently that a subsequently viewed moving grating appears to be motionless while its shape can still be recovered (Cohen, 1965; Hunzelmann & Spillmann, 1984; MacKay, 1982).

Motion Standstill in High Spatial Frequency Gratings

Motion standstill has been extensively studied experimentally with foveally viewed sine-wave gratings and peripherally viewed annular gratings like those illustrated in Figure IV.78-1 (Lyu, 2010). Foveal grating moving with any temporal frequency between 5 to 10 Hz and a spatial frequency of 16 or 22 c/d (it varied between subjects) produced objective and subjective motion standstill nearly all the time for the three subjects formally tested and also for many others. Subjective standstill means that the grating is visible and appears to be standing still. Objective standstill means (a) that the subject can judge the slant of the grating virtually perfectly but (b) cannot report the direction of motion better than chance and (c) that responses to any aspects of the moving stimulus that is perceived as standing still are statistically indistinguishable from responses to interleaved stimuli that are indeed physically motionless. That is, an observer does not distinguish a stimulus that is moving but is perceived as being motionless from a physically motionless stimulus.

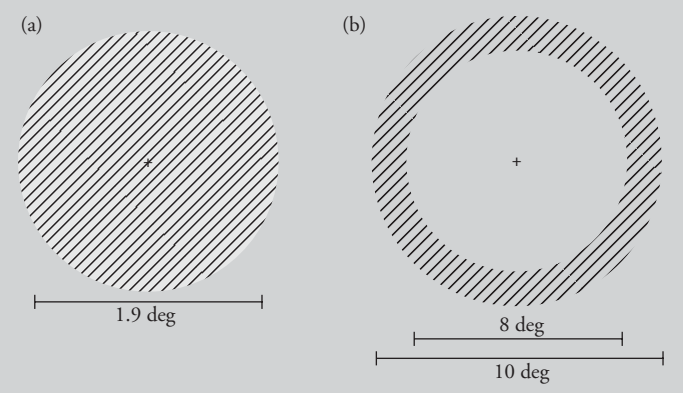


Figure IV.78-1. Snapshots (single frames) of moving sinewave gratings for demonstrating motion standstill in (a) foveal and (b) peripheral viewing. (After Lyu, 2010, with permission.)

The simplest way to experience high spatial frequency standstill is to produce a display (like in Fig. IV.78-1) that has a temporal frequency of about 5 Hz and to view it from various distances. It is usually possible, by moving away from the display, to find a viewing distance at which the slant of the grating is clearly visible but it no longer appears to move. The principle is illustrated in Figure IV.78-2.

Motion Standstill in Equiluminant, Isosalient Chromatic Motion Stimuli

The perception of visual motion is computed by at least three different systems (Lu & Sperling, 1995). The first-order motion system computes motion based on luminance variations. It is exquisitely sensitive and can detect motion with luminance variations of 1 part in 500. Nevertheless, first-order motion cannot be perceived in stimuli that vary in color but are balanced to be precisely equiluminant for the first-order motion system (Lu, Lesmes, & Sperling, 1999a). The stimuli under consideration here (Fig. IV.78-3) were perfectly balanced to reduce first-order contamination to far below threshold (Anstis & Cavanagh, 1983; Lu & Sperling, 2001). Because of the extreme sensitivity of the first-order motion system, the first-order calibration is specific to an observer, to a particular stimulus display, and to the region of the retina in which the stimulus is observed.

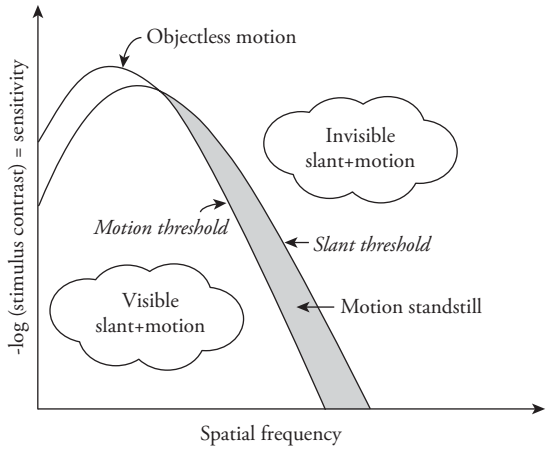


Figure IV.78-2. Schematic illustration of the spatial frequency and contrast that could produce motion standstill for a sinewave grating of a particular temporal frequency, size, and position on the retina. The abscissa represents increasing spatial frequency; the ordinate represents decreasing contrast modulation of the sinewave grating. The line "motion threshold" separates the stimuli for which motion-direction discrimination is possible from those stimuli for which it fails. The line "slant threshold" separates the stimuli for which slant-orientation discrimination is possible from those stimuli for which it fails. Motion standstill occurs in marginal areas where the shape system succeeds and the motion system fails. Objectless motion (not studied here) occurs for stimuli in which contrast is too low to permit shape resolution but nevertheless is sufficient for motion-direction perception. (Figure by George Sperling and Son-Hee Lyu.)

The second-order motion does not respond to luminance directly but to the *variance* of luminance in texture stimuli. It is not relevant to these stimuli because they have insufficient texture.

The third-order motion system responds to salience. Salience in these stimuli is controlled by the relative difference from the background of the red and green regions and by the relative amount of attention they receive from the observer. To achieve the greatest possible contrast, the red output of the display device is set to its maximum value and the green output, which has a greater range, is varied. The green stripes of the grating are then adjusted to isosalience by varying the green saturation, always maintaining equiluminance. When isosalience is achieved, there is no salience variation within the moving grating; therefore there is no salience variation when the grating moves; therefore the third-order motion system has no signal; and therefore no motion is perceived. However, the shape and color systems can function perfectly over a wide range of speeds of the grating, and, over this range, the stimulus will appear motionless to the observer. And of course the observer is absolutely unable to report the direction of motion of the apparently stationary stimulus—it seems to be a stupid question. Motion standstill has been vividly demonstrated for red–green gratings of different spatial and temporal frequencies and for all observers (Lu et al.,1999b).

The first- and third-order calibrations are highly specific to an observer. A change of 0.5% in luminance or 5%

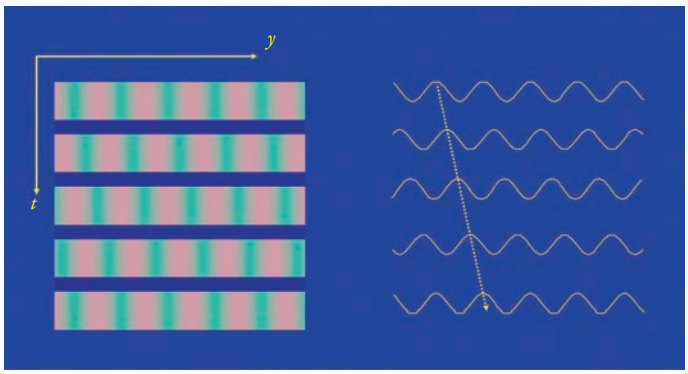


Figure IV.78-3. Five frames of a chromatic sinewave grating in which consecutive frames are translated 90° to the right. When the red and green stripes are calibrated to be equiluminant for a particular observer and then also made to be isosalient (varying the saturation of green so that it appears as different from the background as red) then this five-frame display will appear to be standing still even though nearly everyone else who is looking at it (but not personally calibrated) will see it as obviously moving to the right. (Figure by Zhong-Lin Lu and George Sperling.)

in saturation changes a display from standstill to obvious motion. Therefore, even while the calibrated observer is consistently reporting standstill to a particular display, everyone else (being uncalibrated) looking at the same display is likely to see it as obviously moving, as indeed it is!

Because human color vision is trichromatic, and because only two dimensions of color have been used to produce equiluminance and isosalience (which are each one-dimensional), there is one extra dimension available to produce an enormous range of differently colored equiluminant and isosalient stimuli. Because of technical difficulties, this potential range has hardly been explored.

Stereomotion Standstill

When a dynamic random-dot stereogram as in Figure IV.78-4 produces a horizontal depth grating that translates upward or downward, there are no monocular cues to motion or form and no first-order or second-motion system activation. However, the near plane of the grating tends to be seen as figure (high salience) versus the background (low salience). The moving salience grating is a fine third-order motion stimulus at temporal frequencies of 3 Hz and lower. At 4 to 5 Hz, motion direction discrimination falls to chance, but pattern discrimination (which is one of three possible spatial frequencies) is very accurate, and the perceived speed is mostly judged as zero (i.e., motion standstill; Tseng et al., 2006). Motion standstill in this third-order motion stimulus follows the regime outlined in Fig. IV.78-2 for first-order motion stimuli. Motion standstill in third-order motion system differs from first-order motion standstill mainly in the lower spatial and temporal resolution of the third-order motion system.

Blue-Cone Motion Standstill

Because blue cones make virtually no contribution to luminance, they do not significantly activate the first-order motion system (Herrera et al, 2013). Therefore, blue-cone isolating stimuli offer an opportunity for studying motion standstill in the third-order motion system, where it is much easier to achieve than in the first-order system. However, there has been no systematic study of motion standstill with blue-cone isolating stimuli. Motion slowdown has been studied with stimuli that stimulate mostly but not exclusively blue cones (Cavanagh et al., 1984), and motion standstill has occasionally been observed.

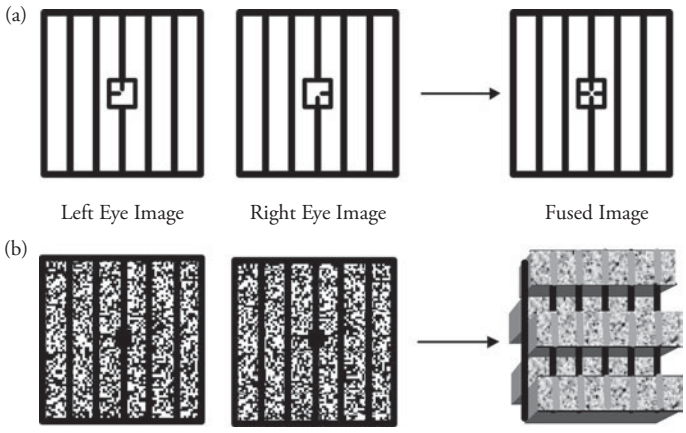


Figure IV.78-4. (a) A stereo stabilizing pattern to insure proper fixation. (b) One frame of a left- and right-eye image of a dynamic random-dot stereogram that produces a raised horizontal grating moving upward or downward. Binocularly free-fusing the left- and right-eye images in (b) produces a stereo perception similar to the representation of the fused image. (After Tseng et al., 2006.)

Standstill Via Motion Adaptation

Cohen (1965) used a variety of stimuli to study peripheral motion adaptation and motion aftereffects. After prolonged viewing between 8.5° to 11° peripherally, linearly translating dots and squares occasionally appeared to stand still. MacKay (1982) reported standstill of a moving extrafoveal grating. Subsequently, Hunzelmann and Spillmann (1984) used slowly rotating black and white disks presented in the periphery to measure the viewing time required for subjects to perceive standstill. The disks appeared to slow down and come to a standstill within 5 to 25 s.

CONCLUSIONS

The motion standstill illusion can be observed in a wide variety of special moving stimuli that either disadvantage or fatigue the motion systems to the point where no motion is perceived but where the shape, texture, color, and depth systems are still able to function sufficiently to extract a stable image from the moving display.

ACKNOWLEDGMENTS

This work was supported in part by National Science Foundation Award BCS-0843897.

REFERENCES

Anstis, S., & Cavanagh, P. (1983). A minimum motion technique for judging equiluminance. In J. D. Mollon & E. T. Sharpe (Eds.), *Colour vision* (pp. 155–166). New York: Academic Press.

Cavanagh, P., Tyler, C. W., & Favreau, O. E. (1984). Perceived velocity of moving chromatic gratings. *Journal of the Optical Society of America A*, *1*(8), 893–899.

Cohen, R. (1965). Adaptation effects and after effects of moving patterns viewed in the periphery of the visual field. *Scandinavian Journal of Psychology*, *6*, 257–264.

Herrera, C., Sun, P., Groulx, K., Wright, C., Chubb, C., & Sperling, G. (2013). How do the S-, M- and L-cones contribute to motion luminance assessed using minimum motion? *Journal of Vision*, *13*(*9*), 1021.

Hunzelmann, N., & Spillmann, L. (1984). Movement adaptation in the peripheral retina. *Vision Research*, *24*(12), 1765–1769.

Lu, Z.-L., Lesmes, L. A., & Sperling, G. (1999a). Perceptual motion standstill in rapidly moving chromatic displays. *Proceedings of the National Academy of Sciences USA*, *96*, 15374–15379.

Lu, Z.-L., Lesmes, L. A., & Sperling, G. (1999b). The mechanism of isoluminant chromatic motion perception. *Proceedings of the National Academy of Sciences USA*, *96*(14), 8289–8294.

Lu, Z-L., & Sperling. G. (1995). The functional architecture of human visual motion perception. *Vision Research*, *35*, 2697–2722.

Lu, Z.-L., & Sperling, G. (2001). Sensitive calibration and measurement procedures based on the amplification principle in motion perception. *Vision Research*, *41*, 2355–2374.

Lyu, S.-H. (2010). Studies of human motion processing: Motion standstill in luminance- and contrast-modulated motion and fMRI brain imaging of stereo third-order motion (Doctoral thesis). University of California, Irvine.

MacKay, D. (1982). Anomalous perception of extrafoveal motion. *Perception*, *11*, 359–360.

Rossi, E. A., & Roorda, A. (2010). The relationship between visual resolution and cone spacing in the human fovea. *Nature Neuroscience*, *12*(2), 156–157.

Sperling, G., Lyu, S.-H., & Kim, H. (2002). Apparent standstill of rapidly moving first-order and second-order motion stimuli [Abstract]. *Perception*, *31*(Suppl.), 100.

Steinman R. M., Cushman W. B., & Martins A. J. (1982). The precision of gaze. *Human Neurobiology*, *1*, 97–109.

Tseng, C.-H., Gobell, J. L., Lu, Z.-L., & Sperling, G. (2006). When motion appears stopped: Stereo motion standstill. *Proceedings of the National Academy of Sciences, USA*, *103*, 14953–14958.

Chapter 79
Objectless Motion

The Pedestalled Motion Paradigm

George Sperling and Zhong-Lin Lu

The clearest case of objectless motion occurs in patients who have suffered brain injuries in which cortical area V1 (visual striate cortex) is so severely damaged that they are blind and have zero pattern or object perception. When presented with motion stimuli, they say they can see nothing, and indeed they cannot report anything greater than chance about stimulus shape or color. Nevertheless, they can report the direction of stimulus motion with accuracies that are significantly greater than chance (e.g., Azzopardi & Hock, 2011). This is possible because there is a neural pathway from the retina to cortical motion area MT that bypasses V1. That is not the case for any other visual modality. In sighted observers, the concept of objectless motion seems to have originated with Wertheimer (1912) in his description of a display that produced what he called phi motion. His phi motion consisted of two adjacent separated objects, flashed alternately with about 30 msec between successive flashes.

Wertheimer's observations were repeated and elaborated recently by Steinman, Pizlo, and Pizlo (2000). In very quick alternations, two adjacent white disks on a dark background appeared stationary and flickering, and pure (objectless) motion occurs between them. That is, there was apparent back-and-forth motion but, at high alternation rates, the only objects (the two disks) were perceived as stationary, merely flickering. Since Wertheimer (1912), there have been several reports of displays to which terms such as "objectless motion" have been applied (Allport, 1968; Ekroll, Faul, & Golz, 2008; Hock & Nichols, 2013; Petersik & McDill, 1981; Saucer, 1953, 1954; Tyler, 1973; Zeeman & Roelofs, 1953). We present here what we consider to be an even more compelling objectless-motion display, the pedestalled motion paradigm. A computational theory exists for this display that both predicts (prior to the demonstration) the magnitude of the motion that is perceived and why it appears to be objectless.

First, we wish to remove from consideration a degenerate case of objectless motion that can be perceived in extremely low-contrast stimuli. In a normally sighted person, it is usually possible to find grating stimuli of extremely low contrast for which motion sensitivity is greater than contrast sensitivity so that motion is perceived when contrast is not. These extremely faint presentations represent objectless motion, but they are hardly impressive illusions.

On the other hand, as in phi motion, in the pedestalled motion paradigm the perception of motion is strong and perfectly reliable, yet every visible object appears to be absolutely stable and motionless, or simply wobbling back and forth, while there is a powerful sensation of linear motion in a consistent direction.

The pedestalled motion paradigm (Lu & Sperling, 1995, 2001) emerged as a prediction (van Santen & Sperling, 1984) from Reichardt's (1961) theory of first-order motion. Figure IV.79-1 illustrates the stimuli. We first consider only the "luminance" stimuli (Figs. IV.79-1d through IV.79-1f) that stimulate primarily the first-order motion system. Figure IV.79-1a schematically illustrates five consecutive frames of a stationary sine wave. Figure IV.79-1d illustrates these same five frames, each of which consists of a cutout midsection of a much larger vertically extended sine-wave grating (Fig. IV.79-2) with the modulation shown in Figure IV.79-1a. When displayed dynamically, the stimulus depicted in Figure IV.79-1d appears as a steady, stationary sine-wave grating called "the pedestal" because of its function in the pedestalled motion paradigm.

Figures IV.79-1b and IV.79-1e illustrate five consecutive frames of a sine-wave grating that moves 90° to the right in consecutive frames, the "moving sine wave." The slanting line indicates the rightward movement of the peak. When the frames of Figure IV.79-1e are shown consecutively at any speed up to about 120 frames/s (30 Hz), they produce a vivid sensation of motion to the right. Critically, the modulation amplitude of the frames in Figures IV.79-1b and IV.79-1e is only half the amplitude of the modulation in Figures IV.79-1a and IV.79-1d. Figures IV.79-1c and IV.79-1f show the algebraic sum of the pedestal and the moving sine wave. In fact, the sum of two sine waves of the same frequency is yet another sine wave of that frequency, although it may differ in phase and amplitude. Because of the greater amplitude of the pedestal, in the summed stimulus, the motion is concealed. There is only a wobble (indicated by the dashed line in Fig. IV.79-1c) in the location of the peaks of the pedestal plus moving sine wave.

When the stimulus in Figure IV.79-1f is shown slowly, for example, at 4 frames/s (one full cycle per second, 1 Hz), the stimulus is perceived as wobbling back and forth, as indicated by the dashed line in Figure IV.79-1c. We know

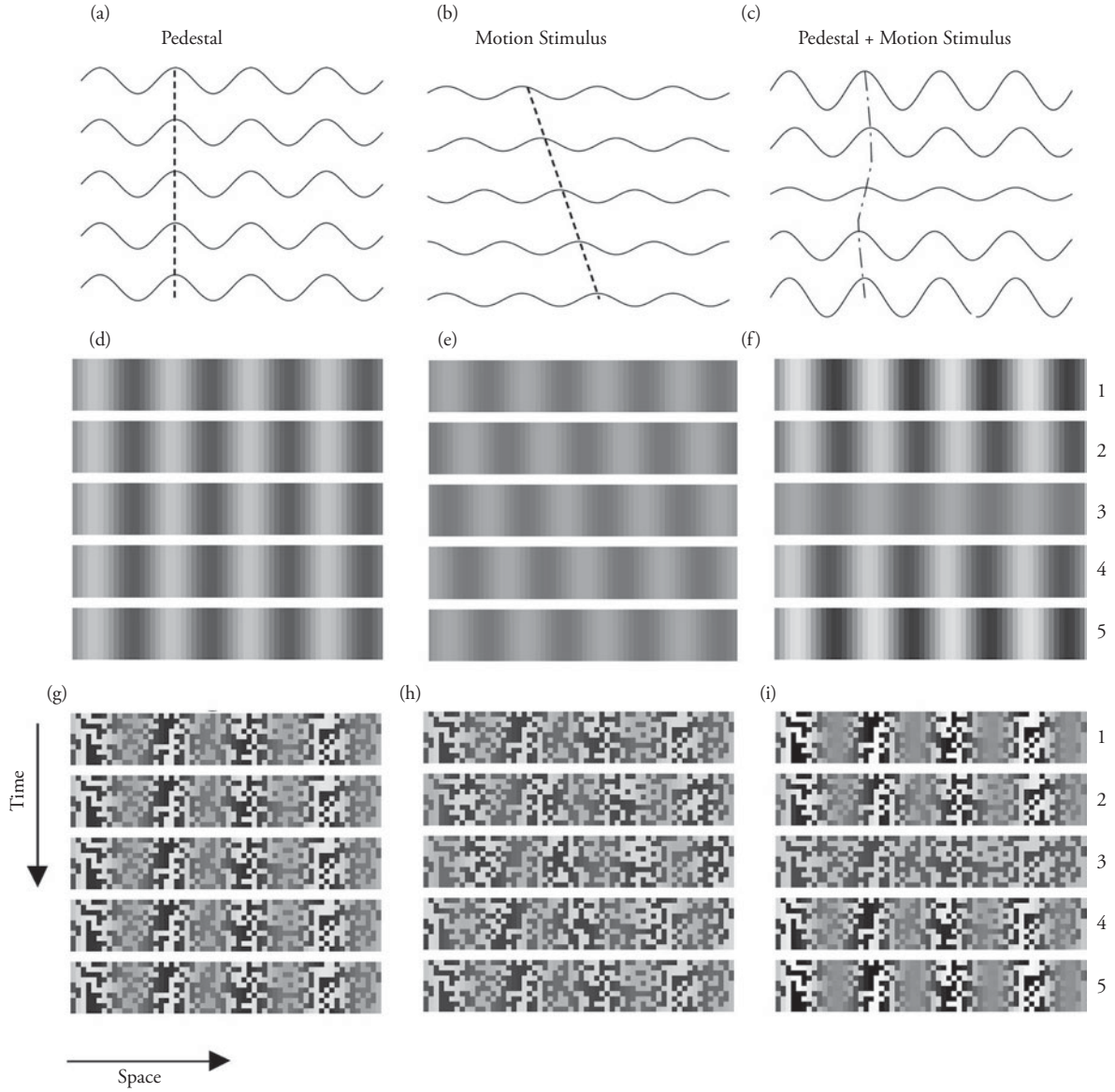


Figure IV.79-1. Five consecutive frames of (a) a stationary sine wave, (b) a moving sine wave, and (c) the sum (a + b). These sine waves are instantiated as luminance-modulated (first-order) gratings in d, e, and f and as texture contrast modulated (second-order) gratings in g, h, and i. Left-to-right motion-direction perception is as good in f, i as in e, h. The pedestalled motion in f, i is perceived as an invisible "wind" above a back-and-forth wobbling grating at low temporal frequencies and above a stationary grating for temporal frequencies greater than about 10 Hz. (After Lu and Sperling, 1995, with permission.)

that at such slow speeds, the perception of motion in such a stimulus is produced primarily by the third-order motion system that tracks the peaks and valleys of the stimulus (Lu & Sperling, 1995). At about 3 Hz, the contribution of the first-order system becomes significant, and the contribution of the third-order motion system begins to diminish. At this point, the real (and perceived) rate of wobble has greatly increased, but there is also a perceived "wind" that seems to flow from left to right above and independently of the wobble. At a temporal frequency of 10 Hz (40 frames/s), for most observers, the wobble is no longer perceived, the grating looks perfectly stationary as in Figure IV.79-2, and only a left-to-right wind seems to pass over the stationary sine-wave grating. For some subjects, the temporal frequency must increase to 12 or 15 Hz before the wobble ceases. Note: We use *temporal frequency* rather than *speed* to describe the movement of a grating because temporal frequency is the critical variable for the motion system.

A remarkable theorem derived from Reichardt's theory (Lu & Sperling 1995; van Santen & Sperling, 1984) is that the strength of a Reichardt model's motion output is just as high in the pedestalled motion display of Figure IV.79-1f as in the ordinary motion of Figure IV.79-1e.[1] Insofar as first-order motion is well described by the Reichardt model (which it is; cf. van Santen & Sperling, 1984), an observer should be just as accurate in judging motion direction in pedestalled motion stimuli as in ordinary motion stimuli. This surprising prediction has been verified numerous times. However, there is a caveat: light adaptation, contrast gain control, and other early visual processes distort



Figure IV.79-2. A low-contrast, stationary sine-wave grating. This is how the pedestalled motion display looks at high temporal frequencies. The perception of left-to-right linear motion (illustrated in Fig. IV.79-1) is not representable pictorially; observers describe it as an invisible "wind" that passes from left to right over the grating. (Figure by George Sperling and Zhong-Lin Lu.)

visual inputs before they reach the stage of motion perception (Lu & Sperling, 1996). Conveniently, low modulation amplitude inputs are distorted less. Therefore, tests of the Reichardt model are performed with stimuli that have modulation amplitudes of less than about 5% (which is still more than 20 times threshold amplitude for most observers). For these stimuli, predictions of the Reichardt model are verified experimentally with mathematical precision (Lu & Sperling, 1995; van Santen & Sperling, 1984).

Pedestalled motion produces as strong a motion perception as ordinary motion. Because the sine-wave pedestal is the only visible object, and because it either merely wobbles or at high temporal frequencies remains motionless, the perceived motion of the moving sine-wave grating does not relate to any visible object and is perceived simply as the direction of an invisible wind—pure, objectless motion. Remarkably, it has been demonstrated repeatedly that observers can report the direction of the wind as accurately in the pedestalled objectless motion display as when the pedestal is absent and they report the direction of motion of a clearly visible moving grating. That observers can clearly see a grating is the result of a shape/object computation; this computation neither helps nor hinders the first-order motion computation, which is independent of the static shape.

Figures IV.79-1g, IV.79-1h, and IV.79-1i show texture contrast modulated (second-order) stimuli analogous to the luminance modulated (first-order) stimuli of Figures IV.79-1d, IV.79-1e, and IV.79-1f. In the second-order stimuli, texture contrast is modulated between high contrast and low contrast analogous to luminance modulation between light and dark. Properly constructed, there is no luminance modulation in the second-order stimuli—their motion is invisible to the first-order system. Nevertheless, these second-order stimuli exhibit the same pedestal immunity as first-order stimuli, and observers now perceive a directional motion "wind" passing over a wobbling or stationary contrast-modulated texture. The wind is second-order objectless motion.

Note that at temporal frequencies of about 3 Hz, the pedestalled second-order stimulus is perceived both to wobble and to have a left-to-right wind. As the first-order motion system has been cancelled, and the wind is second-order motion, the wobble is the result of yet another computation—third-order motion. An interesting property of these multisystem illusions is that, in peripheral viewing, the higher-order system is relatively disadvantaged (Lu & Sperling, 1999). Looking to the side when viewing either first-order or second-order pedestalled motion disadvantages the third-order motion system and thereby diminishes or eliminates the perceived wobble.

In addition to being fascinating illusions, the objectless motion stimuli have two important properties. They demonstrate vividly that the first- and second-order motion computations are separate from and independent of shape and object computations. Also, whereas natural stimuli invariably activate more than one motion system, the pedestalled motions are pure first-order and pure second-order motions; that is, they stimulate only the motion system to which they are directed. This makes pedestal stimuli ideal stimuli for psychophysical and physiological investigations of visual motion systems.

ACKNOWLEDGMENTS

This work was supported in part by National Science Foundation Award BCS-0843897.

NOTE

1. This follows because when two sinusoidal inputs to a Reichardt model have different temporal frequencies (and meet certain other technical conditions ; Lu & Sperling, 2001), the output to the sum of the inputs is equal to the sum of outputs that each would have produced individually. The pedestal is not changing, and therefore it has a temporal frequency of zero. The motion stimulus, by definition, has a nonzero temporal frequency. As the output of a Reichardt model is zero for any stimulus with zero temporal frequency, the motion response is just as strong for a pedestalled as for an ordinary motion stimulus.

REFERENCES

Allport, D. A. (1968). Phenomenal simultaneity and the perceptual moment hypothesis. *The British Journal of Psychology*, *59*, 395–406.

Azzopardi, P., & Hock, H. S. (2011). Illusory motion perception in blindsight. Proceedings of the National Academy of Sciences (USA), 108(2), 876–881.

Ekroll, V., Faul, F., & Golz, J. (2008). Classification of apparent motion percepts based on temporal factors. *Journal of Vision*, *8*(4), 31, 1–22.

Hock, H. S., & Nichols, D. F. (2013). The perception of object versus objectless motion. *Attention, Perception, & Psychophysics*, *75*, 726–737.

Lu, Z.-L., & Sperling, G. (1995). The functional architecture of human visual motion perception. *Vision Research*, *35*, 2697–2722.

Lu, Z.-L., & Sperling, G. (1996). Contrast gain control in first-and second-order motion perception. *Journal of the Optical Society of America A*, *13*(12), 2305–2318.

Lu, Z.-L., & Sperling, G. (1999). Second-order reversed phi. *Perception & Psychophysics*, *61*(6), 1075–1088.

Lu, Z.-L., & Sperling, G. (2001). Three-systems theory of human visual motion perception: Review and update. *Journal of the Optical Society of America A*, *18*, 2331–2370.

Petersik, J. T., & McDill, M. (1981). A new bistable motion illusion based upon "kinetic optical occlusion." *Perception*, *10*, 563–572.

Reichardt, W. (1961). Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In W. A. Rosenblish (Ed.), *Sensory communication* (pp. 303–317). New York: Wiley.

Saucer, R. T. (1953). The nature of perceptual processes. *Science*, *117*, 556–558.

Saucer, R. T. (1954). Processes of motion perception. *Science*, *120*, 806–807.

Steinman, R. M., Pizlo, Z., & Pizlo, F. P. (2000). Phi is not beta, and why Wertheimer's discovery launched the Gestalt revolution. *Vision Research*, *40*, 2257–2264.

Tyler, C. W. (1973). Temporal characteristics in apparent movement: Omega movement vs. phi movement. *Quarterly Journal of Experimental Psychology*, *25*, 171–181.

van Santen, J. P. H., & Sperling, G. (1984). Temporal covariance model of human motion perception. *Journal of the Optical Society of America A*, *1*, 451–473.

Wertheimer, M. (1912). Experimental studies of perceived movement (Experimentelle studien über das sehen von bewegung). *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, *61*, 161–265.

Zeeman, W. P. C., & Roelofs, C. O. (1953). Some aspects of apparent motion. *Acta Psychologica*, *9*, 159–181.

Chapter 80
Silencing the Awareness of Change

Jordan W. Suchow and George A. Alvarez

INTRODUCTION

Determining whether an object is changing is usually a simple matter of looking at it. But sometimes looking is not enough. For example, in the case of *change blindness*, drastic changes to a viewed scene become undetectable either because of a distracting flash that competes for attention or because the changes are slow and gradual (O'Regan, Rensink, & Clark, 1999; Simons, Franconeri, & Reimer, 2000; Simons & Rensink, 2005). In *silencing*, a form of change blindness caused by the presence of motion, rapid and salient changes become hard to notice even without distraction (Suchow & Alvarez, 2011). Here we describe what is known about silencing and its possible causes, starting with the basic phenomenon.

When a set of objects changing in color, brightness, size, or shape moves with respect to the viewer's gaze, it becomes surprisingly hard for the viewer to notice that the objects are changing (Fig. IV.80-1, Video IV.80-1; Suchow & Alvarez, 2011). The discrepancy between the actual and perceived rate of change can be quantified as their ratio—the *silencing factor*. Silencing is a strong effect: when measured in the lab using an adjustment procedure, demonstrations such as Video IV.80-1 typically lead participants to report a perceived rate of change that is one-fourth to one-fifth of the actual rate (Suchow & Alvarez, 2011).

As is the case with many illusions, it is often fruitful to measure how parameters of the display influence the strength of the perceptual effect. A number of parameters are known to influence the strength of silencing:

The faster the motion, the stronger the silencing (Suchow & Alvarez, 2011). The relationship between the strength of silencing and speed (i.e., the angular velocity of a dot as it moves around the ring) is well described by a power law, producing a straight line on a log-log plot (Suchow & Alvarez, 2011). The influence of speed has its limits: if the ring moves very slowly, then its motion is undetectable; if it moves very quickly, then everything blurs. At intermediate speeds, however, when it is possible both to detect the motion and to see the dots, there is a graded effect of speed on the strength of silencing.

The dots in Video IV.80-1 are tightly packed, and the strength of silencing depends on their spacing (Turi & Burr, 2011). Closer spacing produces stronger silencing, independent of the objects' size, and with a critical spacing (i.e., the tightest center-to-center dot spacing for which the awareness of change is spared) that is consistent with the Bouma law of crowding (Bouma, 1970; Pelli & Tillman, 2008; Turi & Burr, 2011).

The initial brightness of each dot in Video IV.80-1 was chosen uniformly over the full range and independently of the rest, such that the whole ring is of varied brightness. Silencing depends on there being heterogeneity among the dots. In fact, it is much easier to notice changes to a set of dots

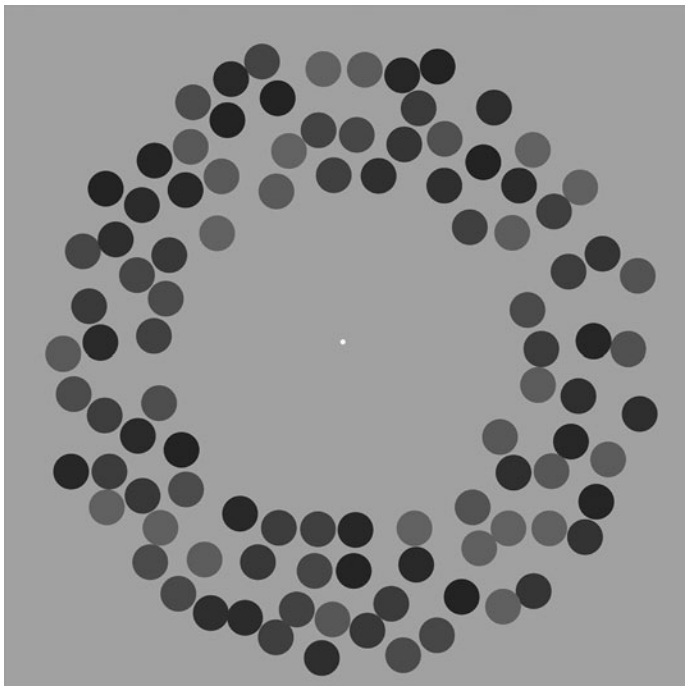
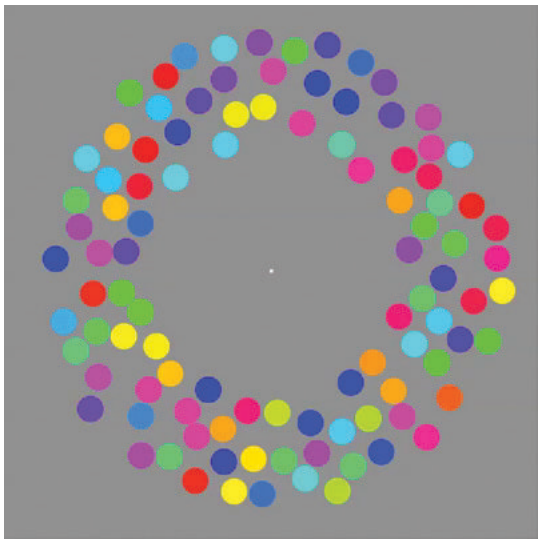


Figure IV.80-1. Motion silences awareness of changes in brightness. One hundred dots are arranged in a ring around a central fixation mark. In Video IV.80-1, from which this single frame has been drawn, each dot changes rapidly in brightness, going from light to dark shades of grey and then back. The initial brightness of each dot is chosen randomly and independently of the rest, such that the whole ring is of mixed brightness. When the ring is motionless, it is obvious to the viewer that the dots are changing. But when the ring rotates about its center (or, in the general case, when it moves with respect to the viewer's gaze), it becomes difficult to tell whether the dots are changing at all. This attenuation in the salience of change, brought on by motion, is the silencing illusion. (Credit: Jordan W. Suchow.)



Video IV.80-1. Motion silences awareness of changes in hue. See the Companion Website (Credit: Jordan W. Suchow).

when those changes produce global shifts in the set's composition (Alvarez & Oliva, 2008; Saiki & Holcombe, 2012). Changes that do not shift the set's composition are hard to notice (Alvarez & Oliva, 2008; Saiki & Holcombe, 2012).

Last, silencing is produced by some kinds of motion but not by others. When the viewer's gaze is held steady as an object moves on the screen, as is the case in Video IV.80-1, two kinds of motion occur at the same time: motion of the object in the world and motion of the object's image as projected onto the retina of the eye. To determine which of these two types of motion is responsible for silencing, it is necessary to dissociate them, which can be accomplished with a simple trick. When we move our eyes to track a moving object, the object moves in the world even though its image is stationary on the eye's retina. In contrast, when we move our eyes while all of the objects in a scene remain stationary, their image moves across the retina even though they are stationary in the world. These two manipulations dissociate the two kinds of motion, making it possible to determine the contribution of each. As it turns out, only motion of the objects with respect to the eyes produces silencing. If the eyes are still and the objects move, there is silencing. If the objects are still and the eyes move, again there is silencing. But if the objects move and the eyes follow them, silencing is markedly absent: the changes are still noticeable (Suchow & Alvarez, 2011). Motion of the objects in the world is irrelevant; all that matters is whether the objects move with respect to the viewer's gaze.

So far we have seen that silencing is modulated by speed, depends on the spacing of the objects, and requires motion across the retina. Together these findings suggest a possible mechanistic explanation of silencing based on known properties of the visual system.

THE BRIEF GLIMPSE HYPOTHESIS (AND ITS DOWNFALL)

Much of the visual system is organized into retinotopic maps, for which neighboring points on the map correspond to neighboring locations in the visual field (Wandell, Brewer, & Dougherty, 2005). Retinotopic organization is ubiquitous in the brain, showing up both in early visual areas like the lateral geniculate nucleus and V1 and in high-level visual areas like the lateral occipital complex and frontal eye fields (Larsson & Heeger, 2006; Wandell et al., 2005). For a moment, let us consider the possibility that the units responsible for detecting change are organized retinotopically.

Figure IV.80-2 shows the implications of retinotopically organized ("local") detectors for the detection of change. Each local detector monitors a fixed region in the visual field, known as its *receptive field*. Because the receptive field of a local detector covers only a small portion of the full visual field, it means that as an object moves from one location to the next, it leaves the bounds of one detector and becomes situated within the reach of another. Thus each detector is afforded only a glimpse of the changing object. The faster the object moves, the briefer this glimpse. Perhaps it becomes too brief to detect the change.

To recover the changes, the viewer needs a linking mechanism that can combine information across multiple detectors. Any such linking mechanism needs to be flexible enough to handle the variety of paths that an object might reasonably take as it travels through the environment. A candidate linking mechanism with the necessary flexibility is attention (Cavanagh, Holcombe, & Chou, 2008; Shioiri, Ogawa, Matsubara, & Yaguchi, 2004). If attention allows for combining information across multiple detectors, then by using attention to track an item as it moves across the visual field, it should be possible to recover some of the changes lost due to silencing.

Whether attention relieves silencing is currently unknown, and published views on it are based solely on the phenomenology of watching the demonstrations (Burr, 2011). However, using the following method, readers can test it out for themselves. First, watch a motionless segment of Video IV.80-1, comparing what happens when you pay attention to all the dots to what happens when you pay attention to just one of them. On a scale of 1 to 10, how much more noticeable does the change become when you attend to one dot? This provides a baseline measure of the effect of attention. Next, watch a moving segment

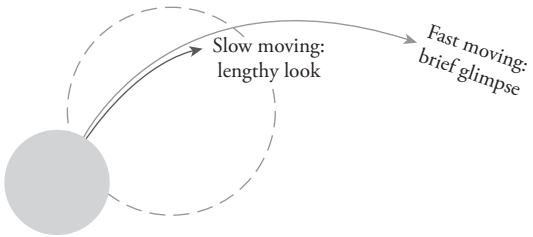


Figure IV.80-2. The brief window hypothesis. The green dot represents a changing object that moves through the visual field. The dotted circle is the receptive field of a hypothetical change detector. When the object moves slowly (dark grey arrow), it spends a considerable amount of time within the detector's receptive field, but when it moves quickly (light grey arrow), it spends less time within the receptive field. Since the amount of time that an object spends within the receptive field is proportional to the object's speed, the glimpse afforded to the detector (and thus the visibility of the change) also depends on speed. (Credit: Jordan W. Suchow.)

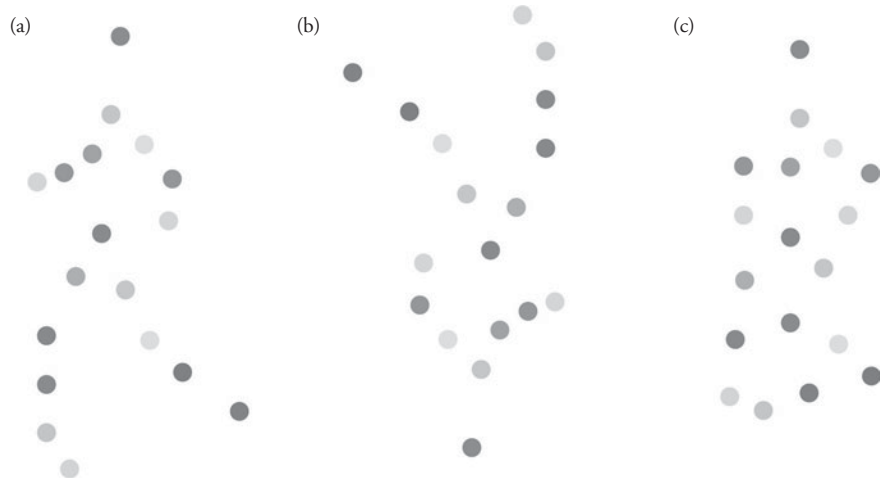


Figure IV.80-3. Perceptual grouping silences awareness of the individual items that comprise the group. In the original demonstration of silencing, the dots are arranged into a ring. Does grouping matter? Yes (Poljac et al., 2012). When the dots are arranged into a familiar figure (A, a point-light walker), the silencing factor is higher than when that same configuration is inverted, becoming unfamiliar and less strongly grouped (B, an inverted point-light walker; Poljac et al., 2012). When the motion is phase-scrambled, such that it no longer resembles the motion of a biological entity, and no longer groups at all, the silencing factor is even lower. This dependence on perceptual grouping is unexplained by the brief glimpse account of silencing. (Credit: Jordan W. Suchow.)

of Video IV.80-1, again comparing what happens when attending to all the dots versus just one of them. On a scale of 1 to 10, how much more noticeable does the change become when you attend to one dot? Comparing your answers for the motionless and moving conditions provides a subjective measure of the influence of attention on silencing. If you rated the effect of attention as being higher in the moving condition than in the motionless one, then your perceptual experience supports the hypothesis that attention relieves silencing. Though the brief glimpse hypothesis is a sensible account of silencing, there is some data that it has trouble accounting for.

The brief glimpse hypothesis cares only about the trajectory of an individual dot and its spacing with respect to the others, not the arrangement or grouping of all the dots. Under this theory of silencing, two displays with the same individual item trajectories, and spacing should produce comparable silencing. In contrast to this prediction, grouping has been found to influence the strength of silencing in the case of point-light walkers (Fig. IV.80-3; Poljac, de-Wit, & Wagemans, 2012). A point-light walker is a set of moving dots whose path is drawn from a digital model of a human walker recorded using motion capture. When each of the dots that comprise the point-light walker change continuously in color, the walker appears to change more when the walker is upright than when it is inverted (Poljac et al., 2012). It is possible that other grouping principles, such as the rigid-body motion versus nonrigid motion, might also impact the strength of silencing.

We also present here three new demonstrations of silencing where the objects do not move much (if at all) in the visual field: motion confined to a short path, illusory motion, and background motion (Fig. IV.80-4, Videos IV.80-2, IV.80-3, IV.80-4). When the moving objects are confined to a short path, each object hovers in the vicinity of a detector, never leaving its sight. When there is only background or illusory motion, the objects do not move at all. The brief glimpse account of silencing predicts that there should be no silencing in these demonstrations. Readers can see for themselves whether these demonstrations lead to silencing.

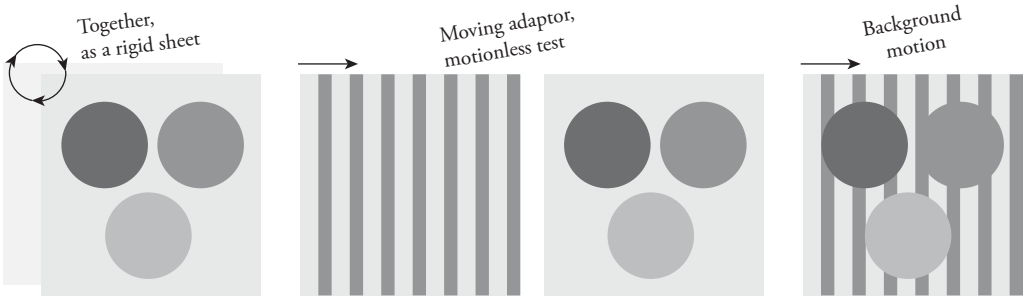
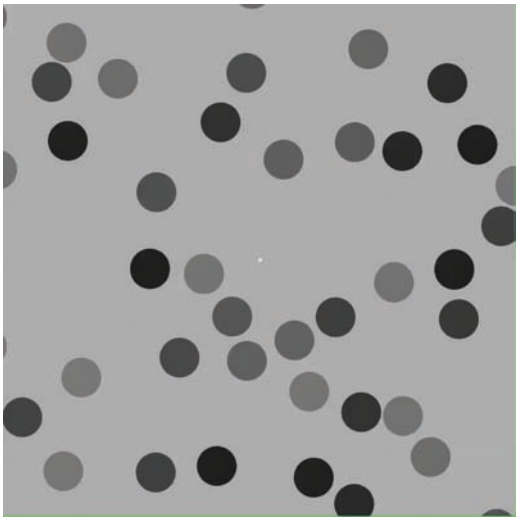
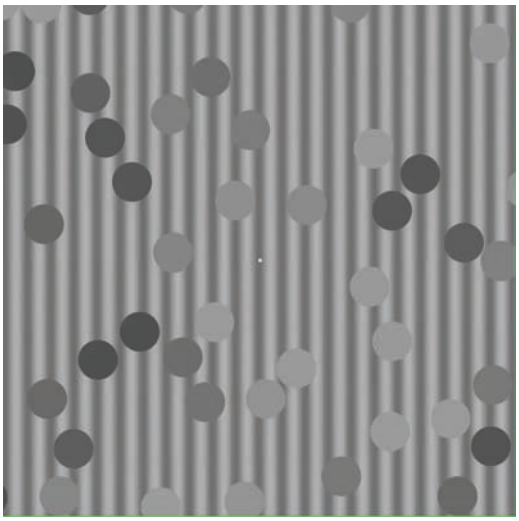


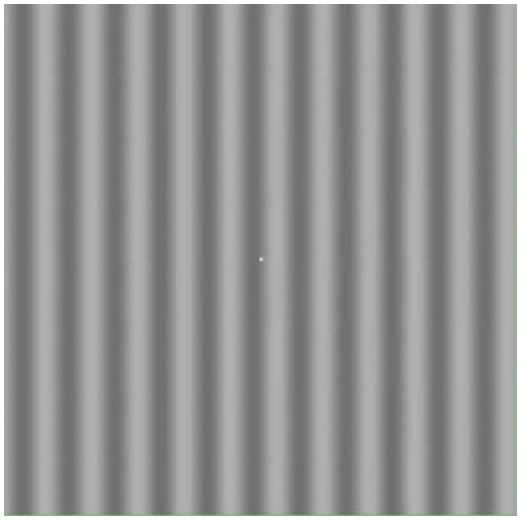
Figure IV.80-4. Three demonstrations of silencing without considerable motion of the objects in the visual field. In the first demonstration (Video IV.80-2), the dots move together as a rigid sheet. In the second demonstration (Video IV.80-3), the viewer adapts to a moving grating before viewing a motionless (but changing) stimulus. The moving grating produces a motion aftereffect in the opposite direction, and this illusory motion then leads to silencing. In the third demonstration (Video IV.80-4), there is background motion: a grating moves in the background while the changing dots remain motionless. (Credit: Jordan W. Suchow.)



Video IV.80-2. Local motion silences awareness of changes in luminance. See the Companion Website (Credit: Jordan W. Suchow.)



Video IV.80-4. Background motion silences awareness of changes in luminance. See the Companion Website (Credit: Jordan W. Suchow.)



Video IV.80-3. Illusory motion silences awareness of changes in luminance. See the Companion Website (Credit: Jordan W. Suchow.)

SILENCING IN HEAD-UP DISPLAYS

The silencing illusion is not merely a laboratory curiosity. Rather it shows up in a variety of real-world applications, such as when presenting information in a head-up display. A *head-up display* is a visual information display embedded in a transparent surface, such as the windshield of a car or aircraft, eyeglasses, camera viewfinder, or gun sight. By overlaying information directly on the viewed scene, these displays can be used to signal important events and information to the viewer. However, because the displays move with respect to their environment, and because the viewer's gaze moves with respect to the display, dynamic elements of the display will vary in salience depending on the surrounding motion context. Some correction of these displays is needed to combat the influence of silencing.

Revealing the limits of human visual perception, determining their impact on performance in real-world tasks and using this knowledge to find ways of improving technology is common in the field of visual perception; this process has been applied to the problem of driving (Beusmans, Aginsky, Harris, & Rensink, 1995), medical screening for cancer (Evans, Tambouret, Evered, Wilbur, & Wolfe, 2011), and gleaning information from signs (Pelli, 2001), among many others.

CONCLUSIONS

Work on the silencing illusion and related phenomena is part of a broad and active area of research on change blindness. Together these studies characterize the mental machinery that allows people to rapidly detect changes in the environment. Some of the most surprising findings in the field have revealed striking limitations in our change detection abilities. Following this tradition, the silencing illusion reveals a new and severe limitation of the change-detection machinery: it is hindered by motion. The strength of silencing has been shown to depend on the type of the motion and its speed, the objects' spacing and their arrangement, and, perhaps, where the viewer attends. These findings await a unifying theory.

ACKNOWLEDGMENTS

This work was supported by a National Science Foundation CAREER Grant (BCS-0953730) to GAA.

REFERENCES

Alvarez, G. A., & Oliva, A. (2008). The representation of simple ensemble features outside the focus of attention. *Psychological Science*, *19*, 392–398.

Beusmans, J. M., Aginsky, V., Harris, C. L., & Rensink, R. A. (1995). Analyzing situation awareness during wayfinding in a driving simulator. In D. J. Garland and M. R. Endsley (Eds.) *Proceedings of the International Conference on Experimental Analysis and Measurement of Situation Awareness* (pp. 245–251). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*, 177–178.

Burr, D. (2011). Visual perception: More than meets the eye. *Current Biology*, *21*(4), R159–R161.

Cavanagh, P., Holcombe, A. O., & Chou, W. (2008). Mobile computation: Spatiotemporal integration of the properties of objects in motion. *Journal of Vision*, *8*, 1–23.

Evans, K. K., Tambouret, R. H., Evered, A., Wilbur, D. C., &Wolfe, J. M. (2011). Prevalence of abnormalities influences cytologists' error rates in screening for cervical cancer. *Archives of Pathology & Laboratory Medicine*, *135*(12), 1557–1560.

Larsson, J. & Heeger, D. J. (2006). Two retinotopic visual areas in human lateral occipital cortex. *The Journal of Neuroscience*, *26*(51), 13128–13142.

O'Regan, J. K., Rensink, R. A., & Clark, J. J. (1999). Change-blindness as a result of "mudsplashes." *Nature*, *398*, 34.

Pelli, D. G. (2001). 47.1: Invited paper: How we see letters: Implications for making better displays. *SID Symposium Digest of Technical Papers*, *32*(1), 1194–1195.

Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience*, *11*(10), 1129–1135.

Poljac, E., de-Wit, L., & Wagemans, J. (2012). Perceptual wholes can reduce the conscious accessibility of their parts. *Cognition*, *123*(2), 308–312.

Saiki, J. & Holcombe, A. O. (2012). Blindness to a simultaneous change of all elements in a scene, unless there is a change in summary statistics. *Journal of Vision*, *12*(3), 1–11.

Shioiri, S., Ogawa, M., Matsubara, K., & Yaguchi, H. (2004). Effect of attention at high temporal frequencies. *Journal of Vision*, *4*, 501.

Simons, D. J., Franconeri, S. L., & Reimer, R. L. (2000). Change blindness in the absence of a visual disruption. *Perception*, *29*, 1143–1154.

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences*, *9*, 16–20.

Suchow, J. W., & Alvarez, G. A. (2011). Motion silences awareness of visual change. *Current Biology*, *21*, 140–143.

Turi M., & Burr D. C. (2011). "Motion silencing" illusion explained by crowding. *Perception*, *40*, 201.

Wandell, B. A., Brewer, A. A., & Dougherty, R. F. (2005). Visual field map clusters in human cortex. *Philosophical Transactions of the Royal Society of London, B*, *360*, 693–707.

Chapter 81
The Kayahara Silhouette Illusion

Nikolaus F. Troje

THE ILLUSION

In 2003 the computer graphics artist Nobuyuki Kayahara published on his website a collection of animations to demonstrate the power of a new web-based graphics programming tool, Macromedia's Flash. One of the animations showed the spinning silhouette of a young girl in a dancing pose, which he called the silhouette illusion (http://www.procreo.jp/labo/labo13). The silhouette appears in black on a background that consists of gradients of grey. The girl stands on one foot, while the other one is stretched out. The position in which she holds her arms is also asymmetric, with one arm closer to the body and the other bent away. The figure itself does not change its posture but spins rigidly about a vertical axis. Centered about the leg on which the figure stands it is revolving at a speed of 300° per second. While doing so, the whole figure bobs up and down once per revolution. At the lowest position one of the feet seems to touch the floor and at its highest position it is about a foot's width above the floor. The floor is not explicitly depicted but implied by the shadow of the spinning girl. The shading of the background seems to be consistent with a light source that creates the shadow in the middle of a spot-lighted center.

The perceptual conflict that grants Kayahara's animation the status of a visual illusion arises from the fact that the girl can be perceived spinning either clockwise or counterclockwise and that the perceived spinning direction changes spontaneously even though nothing changes in the constantly looped animation itself. The silhouette illusion therefore classifies as a reversible figure.

The silhouette illusion belongs to the class of depth-ambiguous reversible figures, such as the Necker cube (Necker, 1832) or Schröder's staircase (Schröder, 1858, see also Figs. IV.81-1a and IV.81-1b). In these two examples, geometrical regularities (such as parallel lines) constrain the interpretation of the projection and impose a particular structure upon it. What is missing, however, are obvious depth cues that distinguish between two possible versions of that structure, which are identical except for a mirror flip about the image plane. In the silhouette illusion it is the familiarity with the shape and articulation of a human body that suggests a likely three-dimensional (3D) structure. However, like for other depth-ambiguous stimuli, this structure is defined only up to a mirror reversal about the image plane. At least at a first approximation, the likelihoods that either of the two versions created the projected silhouette are about equal, and the a-priori probabilities of the two mirror symmetric versions of the dancer figure being a valid object of the real world are also the same.

Everything said so far applies to any static rendering of a silhouette (Fig. IV.81-1c). However, Kayahara's silhouette is not just a static snapshot but shows the figure rotating in depth. It therefore takes advantage of the "kinetic depth effect" first described under this term and investigated experimentally by Wallach and O'Connell (1953). Even for entirely unfamiliar objects, observers reliably interpret the deforming projection of a rotating object in terms of a 3D rigid object rather than a nonrigid two-dimensional (2D) shape. Assuming rigidity of the 3D object and knowledge about the sequence of viewpoints, the array of silhouettes constrains the shape of the object to lie within the limits of what Laurentini (1994) termed its "visual hull." The visual hull is identical to the shape of the object except in areas with pronounced local surface concavities.

The spinning Kayahara silhouette's rotation provides the observer with enough information to recover the shape of its 3D structure. However, since explicit information about the spinning direction (and therefore the geometry of the viewpoints) is missing, there are still two valid solutions left that differ only by a mirror reversal about the image plane.

This already demonstrates that reversal of perceived depth is inevitably coupled with a number of other changes in the properties of the perceived 3D object. All of them are direct consequences of the mirror reversal about the image plane.

The first is the change in chirality (i.e., in handedness). This change is not directly visible in a bilaterally symmetric object like the Necker cube but becomes obvious for the asymmetric posture of the dancing girl. What has been the girl's left hand in one percept becomes her right hand once it flips to the other percept. The same happens to all other body parts too (Fig. IV.81-2).

The second is the orientation with respect to the observer. If the object, or any of its parts or properties (including the axis of rotation), is slanted with respect to the line of sight of the observer, the slant changes its sign

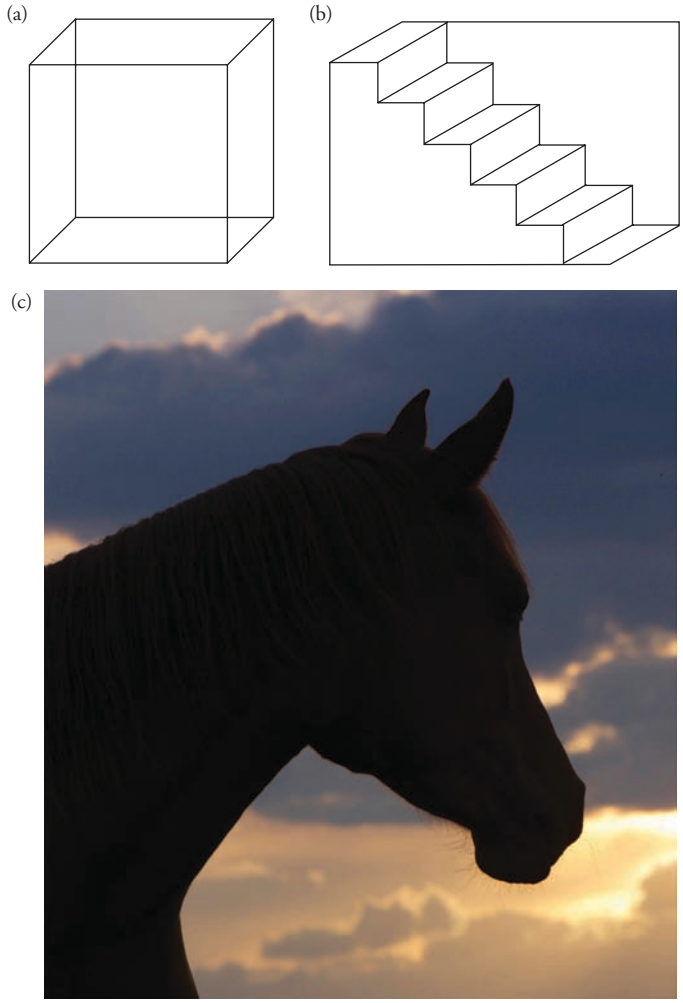


Figure IV.81-1. Three examples for depth-ambiguous figures. (a) The Necker cube is always seen as a cube even though the 2D drawing would also be compatible with many other 3D objects. However, the comparatively high likelihood that lines that are parallel in the projection are generated by lines that are also parallel in the underlying 3D object constrains our perception. Yet even under these constraints, two solutions are possible that differ only in a mirror-flip about the image plane. Both solutions are cubes, but they appear in different orientation. Note that the initial percept seen when first looking at the Necker cube is almost always the viewing-from-above interpretation. (b) Schröder's staircase is another example of a simple depth-ambiguous geometry. (c) The silhouette of the head of this horse can be interpreted in two ways. In one case the horse is oriented toward the observer. The horse's right ear appears in the image on the left side. The brows seen in the silhouette belong to its left eye. In the other percept, the horse is oriented away from the observer. The left of the two ears in the image is in fact the horse's left ear and the brows belong to the right eye. (© Juniors Bildarchiv GmbH/Alamy Stock Photo.)

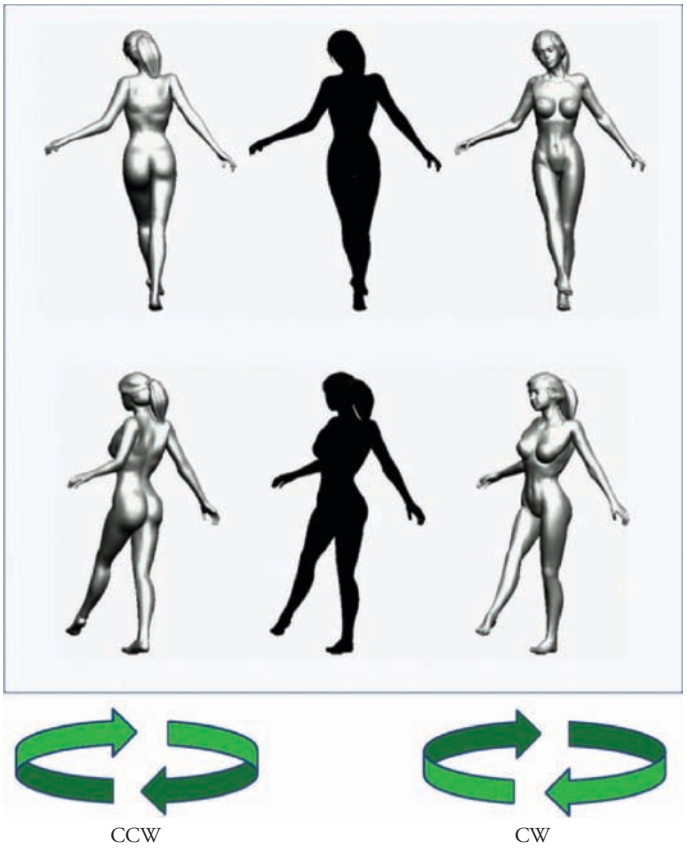


Figure IV.81-2. The silhouette shown here is not identical with Kayahara's but was created to resemble it to a large degree. It is based on a 3D model that was used to illustrate how depth reversal affects the apparent camera elevation, the handedness of the figure, and the perceived direction of rotation. In the left column we see the figure from a slightly lowered viewpoint, the leg on which she stands is the right one, and the direction of rotation from the first to the second frame is counterclockwise. In the right column, the figure appears as seen from a slightly elevated viewpoint; she stands on her left leg and spins clockwise. The silhouettes in the middle column are perfectly consistent with either of the two interpretations. (Adapted from Troje & McAdam, 2010.)

upon depth reversal. This is obvious for the Necker cube of Figure IV.81-1a. In one percept, the cube seems to be slanted toward the observer and is seen as if looked at it from a slightly elevated viewpoint. In the other percept, the cube is slanted away from the observer, who perceives it as if looking up to it from a lower position. The same is also the principle behind Schröder's staircase (Fig. IV.81-1b). In Kayahara's silhouette figure, the vertical axis of the figure around which it rotates and the line of sight of the observer are also not perfectly perpendicular. As a consequence, depth reversal results in a flip between a viewpoint that looks at the figure from a slightly elevated position and a viewpoint that looks at it from slightly below.

The third change, finally, affects the direction in which a rotating object seems to spin. Mirror reversal of the perceived 3D shape means that spinning direction also reverses. Descriptions of the phenomenon itself have a long history that goes back to Sinsteden (1860), who observed that the perceived spinning direction of a windmill seen as a silhouette against the sky is perceptually ambiguous. Later, the phenomenon was termed the "windmill illusion" and investigated by Miles (1931) by back-projecting the shadow of a slowly rotating fan onto a screen.

In Kayahara's animation, in one of the two possible percepts, the girl is standing on her left leg with the right arm bent away from her body. The viewpoint is slightly from above, and the figure spins clockwise. In the other percept, the girl stands on her right leg with the left arm pointing

away from her body, she is seen slightly from below, and she spins counterclockwise.

WHY DOES SHE SPIN CLOCKWISE MORE OFTEN THAN COUNTERCLOCKWISE?

For a number of years, Kayahara's silhouette illusion was enormously popular on the Internet. It was used as an eye-catcher for all sorts of purposes, often promoted as a quick "psychological test" that would reveal to observers whether they were "left-brained" or "right-brained." The idea, which evolved on Internet discussion forums and popular science sites, that perceived spinning direction is related to hemispheric dominance and laterality might have been driven in part by the lack of an explanation for an asymmetry in the frequency of left- and right-spinning percepts: The silhouette illusion seemed to be spinning more often clockwise than counterclockwise. Some popular science blogs collected substantial data about that phenomenon. For instance, an online poll conducted by Cognitive Daily in 2008 reported that about two-thirds of the 1,600 people who answered the poll saw the dancer as spinning clockwise and only one-third as counterclockwise.

Since the original Kayahara illusion was used for that experiment, this also meant that two-thirds of the observers saw the dancer as if looking at it from a slightly elevated viewpoint, and only one-third saw the dancer as if looking at it slightly from below. A study by Troje and McAdam (2010) in which apparent spinning direction was decoupled from the apparent camera elevation showed that the bias observed in Kayahara's silhouette illusion can in fact be fully explained in terms of a preference to see the spinning figure from above rather than from below.

Hints to a preference for perceiving depth-ambiguous stimuli from above rather than from below can be found across the literature (Orbach, Ehrlich, & Helen, 1963; Sundareswara & Schrater, 2008; Washburn, Mallay, & Naylor, 1931) but rarely in a very explicit form. Ooi, Wu, and He (2001) provided experimental evidence that relative height in the visual field is used as a depth cue, which in turn implies the viewing-from-above assumption. Mamassian and Landy (1998) document a viewing-from-above bias for line drawings of surface shapes. Liu and Todd (2004) demonstrate that, at least in some cases, the often-documented light-from-above bias (Ramachandran, 1988) observed when shape is derived from its surface shading is in fact due to a viewing-from-above bias.

THERE IS MORE TO IT

So far we have assumed that the Kayahara illusion does not contain any depth cues at all and that therefore both 3D interpretations are equally likely to underlie the rotating 2D silhouette. It turns out that this is not entirely true. In fact, Kayahara's silhouette does contain two explicit cues that could in principle reveal the "true" depth of the animation. However, the first one is relatively weak, and the second one provides ambiguous information in itself.

The first one is linear perspective. Troje and McAdam's (2010) study reports careful measurements on Kayahara's silhouette figure that revealed the camera parameters used to render the original version of it. They first subtracted the vertical bobbing motion and then traced the ellipse circumscribed by the tip of the toe of the outstretched leg of the dancer. The aspect ratio of this ellipse revealed that the camera used to render the dancer was located either 6.8° above or below the 3D figure of the dancer. Interestingly, when doing the same analysis with the outstretched hand, the camera elevation turned out to be slightly smaller, namely only 6.0°. The measurements indicate that the silhouette was rendered with a camera that introduces some linear perspective. Assuming that the circles circumscribed by foot and hand have a vertical distance of 0.8 m, it can be calculated that the location of the camera was 60 m away from the silhouette. Assuming a height of the whole figure of 1.7 m, the field-of-view angle is 3.2°. Whether or not this rather small degree of linear perspective is perceptible is hard to judge, as thresholds for linear perspective depend strongly on the presence of structural regularities of the object (Freeman, 1966; Saunders & Backus, 2006). The figure does not contain parallel lines or right angles, but the horizontal circles subscribed by features such as the toe and hand that Troje and McAdam (2010) used to infer the camera perspective might as well be used by the human visual system to estimate perspective convergence. Note, however, that the sinusoidal vertical bobbing motion might severely masks this cue.

The second cue, the shadow the figure seems to cast onto the floor, provides conflicting information. As the figure bobs up and down, the distance between the figure and its shadow increases and decreases as expected, supporting the percept that the floor is horizontal. If the floor is horizontal and the shadow is visible below the figure, then we must be looking down onto the floor and therefore also down onto the figure. As explained earlier, the figure should therefore be perceived to rotate clockwise. However, the motion of the shadow also provides a cue, and it points in the other direction. The fact that the tip of the outstretched toe as well as all other features of the shadow rotate counterclockwise in the 2D projection plane indicates that the figure itself rotates counterclockwise too.

The shadow is therefore an impossible shadow and, in a way, Kayahara's silhouette figure not only falls into the category of reversible figure but is also an impossible one. Like the Penrose Triangle (Penrose & Penrose, 1958) or some of M. C. Escher's works (Locher, 1981), it presents a geometry that is locally consistent but globally very unlikely, if not impossible.

These inconsistencies might very well contribute to the appeal of the Kayahara illusion. They are subtle enough not to be noticed easily by the observer, but they probably puzzle the visual system nevertheless, creating the sophisticated tension between expectations and violations of them that often characterizes appealing artwork and that might be a central force that determines an observer's appraisal of it. No matter whether these additional ambiguities were intended by the artist, or whether they are an expression of his creative intuition, they provide very good reasons why Kayahara's silhouette illusion attracts attention and has become such a popular and widely distributed phenomenon on the Internet.

REFERENCES

Cognitive Daily (2008) https://web.archive.org/web/20160624062429/https://scienceblogs.com/cognitivedaily/2008/10/10/casual-fridays-tk421-why-cant/

Freeman, R. B. Jr. (1966). Absolute threshold for visual slant: The effect of stimulus size and retinal perspective. *Journal of Experimental Psychology*, *71*(2), 170–176.

Laurentini, A. (1994). The visual hull concept for silhouette-based image understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *16*(2), 150–162.

Liu, B., & Todd, J. T. (2004). Perceptual biases in the interpretation of 3D shape from shading. *Vision Research*, *44*(18), 2135–2145.

Locher, J. L. (Ed.). (1981). *M. C. Escher: His life and complete graphic work* Amsterdam: Harry N. Abrams.

Mamassian, P., & Landy, M. (1998). Observer biases in the 3D interpretation of line drawings. *Vision Research*, *38*(18), 2817–2832.

Miles, W. (1931). Movement interpretations of the silhouette of a revolving fan. *The American Journal of Psychology*, *43*(3), 392–405.

Necker, L. A. (1832). Observations on some remarkable phenomena seen in Switzerland; and an optical phenomenon which occurs on viewing of a crystal or geometrical solid. *Philosophical Magazine*, *1*, 329–343.

Ooi, T., Wu, B., & He, Z. (2001). Distance determined by the angular declination below the horizon. *Nature*, *414*(6860), 197–200.

Orbach, J., Ehrlich, D., & Helen, A. H. (1963). Reversability of the Necker cube: I. An examination of the concept of "satiation of orientation." *Perceptual and Motor Skills*, *17*(2), 439–458.

Penrose L. S. & Penrose R. (1958) Impossible objects: a special type of visual illusion *British Journal of Psychology* 49, 31–3

Ramachandran, V. (1988). Perception of shape from shading. *Nature*, *331*, 163–166.

Saunders, J. A., & Backus, B. T. (2006). The accuracy and reliability of perceived depth from linear perspective as a function of image size. *Journal of Vision*, *6*(9), 933–954.

Schröder, H. (1858). Über eine optische Inversion bei Betrachtung verkehrter, durch optische Vorrichtung entworfener, physischer Bilder. *Annalen der Physik und Chemie*, *181*, 298–311.

Sinsteden, W. J. (1860). Über ein neues pseudoskopisches Bewegungphänomen. *Annalen der Physik und Chemie*, *187*, 336–339.

Sundareswara, R., & Schrater, P. R. (2008). Perceptual multistability predicted by search model for Bayesian decisions. *Journal of Vision*, *8*(5), 12, 11–19.

Troje, N. F., & McAdam, M. (2010). The viewing-from-above bias and the silhouette illusion. *i-Perception*, *1*, 143–148.

Wallach, H., & O'Connell, D. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, *45*(4), 205–217.

Washburn, M., Mallay, H., & Naylor, A. (1931). The influence of the size of an outline cube on the fluctuations of its perspective. *The American Journal of Psychology*, *43*(3), 484–489.

Chapter 82
The Motion Aftereffect

Frans A. J. Verstraten and Peter J. Bex

THE PHENOMENON

The motion aftereffect (MAE), also known as the movement aftereffect or the aftereffect of seen movement, undoubtedly belongs to the best known of the visual illusions. It occurs after one has looked at a moving pattern for some time. Typically, illusory motion is perceived in the opposite direction at the time the pattern comes to a sudden stop or when the eyes are diverted to a stationary part of the scene, like the rocks next to a waterfall (Anstis, Verstraten, & Mather, 1998; Mather, Verstraten, & Anstis, 1998).

HISTORY

To our knowledge, the first to describe the MAE was Aristotle (384 BC–322 BC) in the treatise "On Dreams" in his series *Parva Naturalia* (in Ross, 1955). Aristotle, as well as Lucretius (50 BC/1994), experienced the illusory motion after having looked at a streaming river. There was some discussion as to who was the first to report the direction of the illusory motion (Gregory, 1981), but so far the credit seems to go to Lucretius (Verstraten, 1996). More modern reports came from Purkyně (1820) and Addams (1834). Purkyně experienced illusory motion after watching a cavalry parade, while Addams saw the rocks next to the lower fall of Foyers moving upward after have fixated on a rock in the center of the waterfall for some time (see Fig. IV.82-1). Addams's description became classic in vision science, and often the MAE is referred to as the waterfall illusion (see Wade & Verstraten, 1998).

MODERN STIMULI

There are several ways to invoke MAEs in a laboratory setting. Bowditch and Hall (1881) used a grating of black and white stripes drawn on a vertical conveyor belt, drifting behind an aperture. Among the most effective stimuli for generating aftereffects are spirals, whose aftereffects are also referred to as the spiral aftereffect. If one does not take Wohlgemuth's monograph in 1911 into account, the first book on MAE was by Holland in 1965 with the title *The Spiral Aftereffect*. Nowadays, computer displays have replaced rotating discs and paper-and-scissor stimuli; however, gratings, as well as random dot patterns, remain popular for invoking MAEs.

NEURAL SUBSTRATE AND DEFINING THE MOTION AFTEREFFECT

In a now-classic paper, Barlow and Hill (1963) seemed to have found the perfect neural correlate for the illusion. They stimulated a rabbit's retina with movement and recorded from retinal ganglion cells. While at rest, these cells typically show spontaneous baseline activity. As soon as a moving stimulus was presented, the firing rate of the cells increased. In time, the firing rated gradually decreased somewhat but stayed well above baseline. When the movement was stopped, the firing rate immediately fell below baseline. This was followed by a recovery from adaptation period over several seconds.

In order to understand the direction of the MAE, one must assume a coupling of at least two populations of motion sensors tuned for opposite directions. This idea was nicely illustrated by John Frisby (1979) and is reproduced in Figure IV.82-3.

This model is known as the ratio model. However, several studies have shown that the MAE is not always opposite to the adaptation direction. For example, in the case of transparent motion consisting of two patterns moving in orthogonal directions, the MAE will normally be unidirectional. That is, although two clear motion directions are perceived during adaptation, the subsequently perceived MAE has only one direction. The MAE direction is opposite of a motion-sensitivity weighted vector sum of the individual transparent patterns (Verstraten, Fredericksen, & van de Grind, 1994). The same is true if inducing patterns that move in two different directions are alternated in time during the adaptation period (Riggs & Day, 1980). One alternative model that is generally more accepted is known as the distribution shift model, in which all the population motion sensors, coding for all directions, are taken into account, not just those that are oppositely tuned (Mather, 1980; Sutherland, 1961).

This period in which neural substrates recover from adaptation to baseline activity is often considered the neural correlate of the perceived illusory motion. However, whether that hypothesis is correct has yet to be proven. Not perceiving



Figure IV.82-1. The lower falls of Foyers. This waterfall is located close to Loch Ness in Scotland. Loch Ness was named after this waterfall. Loch Ness translates to "lake at the waterfall." Fall of Foyers From "Scotland Illustrated." Artist: Thomas Allom Engraver: E. Radclyffe. Publisher: George Virtue London Circa: 1840.

the aftereffect does not mean the aftereffect is not there. That is, part of the recovery from adaptation is not necessarily accompanied by visual (visible) illusory motion, and adaptation can still be present although below the perceptual threshold. Therefore the illusory motion is not simply the correlate of neurons recovering from adaptation until they return to baseline activity levels (Verstraten, Paffen, Carlson, van der Smagt, & Hogendoorn, 2007).

MEASURING THE MOTION AFTEREFFECT

The definition of the MAE is related to the method of its measurement. The classic way to quantify the MAE is to measure its duration by the use of a stopwatch. This measure, however, clearly relies on the visible period of the recovery from adaptation. This method has been used often and successfully (see also Pantle, 1998).

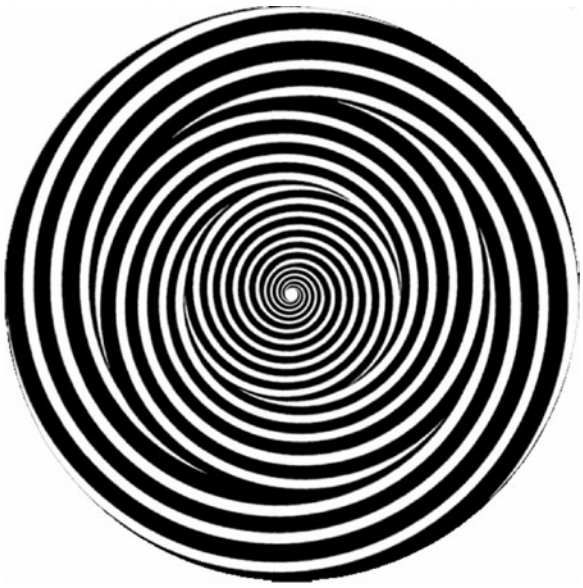


Figure IV.82-2. A very potent stimulus to invoke a MAE is a spiral.

Another measure is known as nulling. It assumes that the illusory movement in the aftereffect could be nulled by real motion in the opposite direction. That is, if the MAE appears to move upward, one could add an equivalent quantity of real motion in the downward direction such that the test would appear to be stationary.

A modern nulling method was introduced in the early 1990s (e.g., Hiris & Blake, 1992). It is based on so-called motion coherence stimuli, composed of moving dots with different levels of signal and random directions. The percentage of dots moving in a signal direction is manipulated. For example, a coherence level of 0% means that all dots are moving in random directions, while 100% coherence means that all dots are moving in a signal direction (Newsome & Paré, 1988). The strength of the MAE is determined and quantified by the percentage of dots moving in the opposite direction of the MAE that are needed to make the stimulus appear nondirectional.

THE PSYCHOPHYSICIST'S MICROELECTRODE?

John Mollon (1974, p. 479) wrote: "If you can adapt it, it's there" meaning that if one finds aftereffects, neural substrates must exist that underlie this phenomenon. For that reason, after effects have been called the psychophysicist's microelectrode. In essence it was seen as a tool to investigate the organization of the motion-sensitive areas of the brain. This "tool" has been quite successful. Next we discuss some of the main findings.

INTEROCULAR TRANSFER, MONOCULAR, AND BINOCULAR SYSTEMS

One technique that has been used to study the organization of the visual system is known as interocular transfer. If one shows adapting motion to one eye while the other eye is

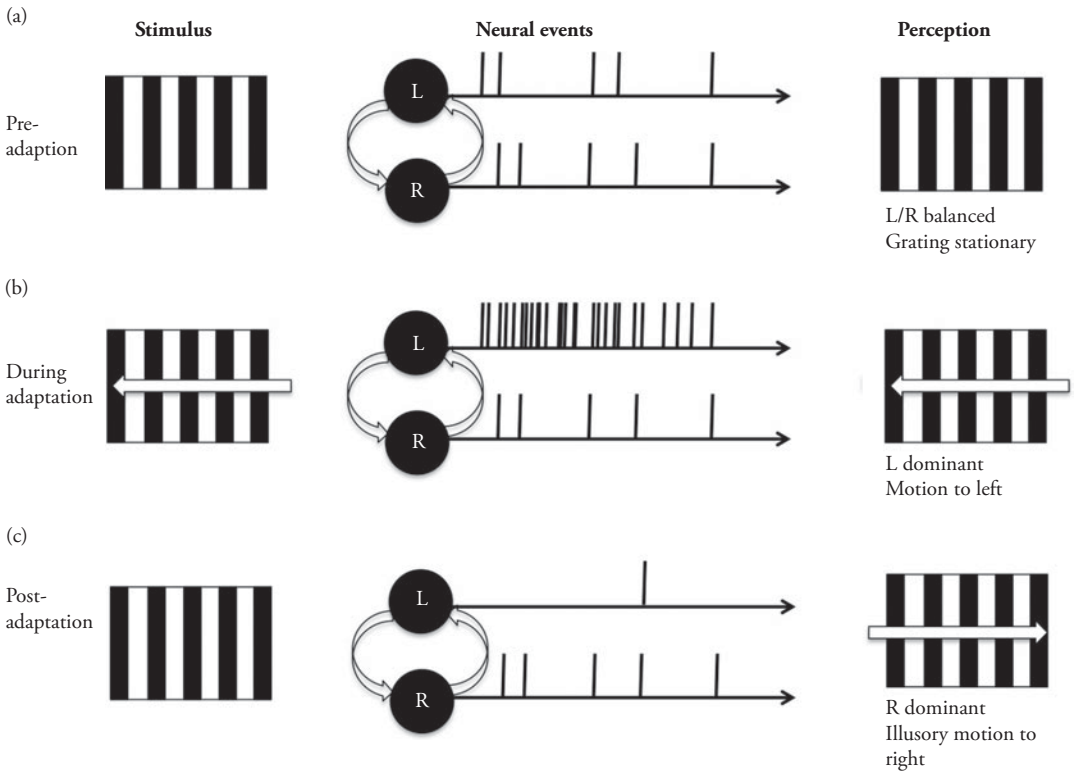


Figure IV.82-3. A representation of the ratio-model explanation for MAEs: Initially, during preadaptation, there is no motion information; the motion sensors tuned for left- and rightward motion are at baseline activity. The activity is balanced between both directions, and as a result the observer will perceive a stationary grating (a). As soon as the grating starts moving to the left, the neurons coding for this direction increase their activity and motion to the left is perceived (b). Once the motion stops, the activity of sensors coding for leftward motion drops below baseline and therefore the sensors for the opposite direction are at a relatively higher level of activity. As a result motion to the right is perceived (c). (Adapted from Frisby, 1979.)

closed, it is possible to perceive illusory motion when a test pattern is presented to the previously closed eye while the previously stimulated eye is closed. This transferred aftereffect generally does not last as long as when the adapted eye is tested and is greatly reduced or absent in people with impaired binocular vision (Hess, Demanins, & Bex, 1997). The interocular transfer is evidence for the involvement of binocular cortical cells, that is, cells that receive input from both eyes, which exist only in cortical areas subsequent to layer 4B of primary visual cortex in mammalian brains (Hubel & Wiesel, 1962).

The interaction between monocular and binocular systems depends on how the adapting and test stimuli are presented. Anstis and Moulden (1970) designed a stimulus for each eye that, when viewed by each eye on its own, was rotating clockwise. However, when viewed with both eyes, the fused stimulus was rotating counterclockwise. After adaptation, monocular testing showed aftereffects based on the monocular retinal direction (a counterclockwise aftereffect), whereas binocular testing showed aftereffects based on the cyclopean perceived direction (a clockwise aftereffect). Thus the MAE originates at multiple levels and can be used to probe the organization of the binocular visual system.

TUNING AND NEURAL CODING

The MAE has provided insight into the computational processes of the visual system. The magnitude of the MAE depends on the spatial and temporal frequencies of adapting and test stimuli. In general, the magnitude of the MAE increases with the spatial similarity between adapting and test stimuli (Over, Broerse, Crassini, & Lovegrove, 1973). However, the largest MAEs are reported for low temporal frequency test patterns, regardless of the temporal frequency of the adapting stimuli (Bex, Verstraten, & Mareschal, 1996). This pattern of results is consistent with motion sensors that are selective for a narrow range of spatial frequencies but a broad range of temporal frequencies, with a peak around 1 Hz.

ADAPTED MOTION SENSORS?

It is tempting to conclude that the MAE is the direct consequence of the attenuation of activity of adapted motion sensors. However, there are several reasons to believe that the story is more complicated: for example, full field stimulation does not result in a visible aftereffect. If one could get close enough to Niagara Falls so that the visual field were completely filled with the falling water, one would experience vection (like when sitting in a train that appears to leave the platform when actually the neighboring train starts moving, as discussed elsewhere in this book) but not a MAE (e.g., Aitken, 1878; Wohlgemuth, 1911). It is hard to argue that motion sensors were not stimulated under these conditions.

Another piece of evidence against activity-based models is the so-called phantom MAE. In this case, moving stimuli

are presented only in some parts of the screen, while other parts are blank. A MAE can be seen and/or measured at locations on the display that do not contain any motion during adaptation (Weisstein, Maguire, & Berbaum, 1977). Furthermore, the direction of phantom MAEs can follow global patterns of optic flow, such as expansion or contraction, even though local induced directions are simple translations, such as leftward or rightward motion (Snowden & Milne, 1996).

Related are aftereffects of relative motion, which also challenge simple models based on sensory response attenuation. A stationary grating will appear to move when it is flanked by moving gratings. This is known as induced motion (as discussed elsewhere in this book). Once the flankers are stopped, these stationary parts will show illusory motion. Again, the local motion sensors at the spatial location where an aftereffect is perceived are not stimulated by real motion (Anstis & Reinhardt-Rutland, 1976). A variation of this experiment was performed by Swanston and Wade (1992). The results also show that MAEs are based on the relative outputs of populations of motion detectors rather than simple changes in the activity of individual motion sensors.

ATTENTIONAL MODULATION

The modulatory effects of attention also show the involvement of higher-order perceptual processes. Chaudhuri (1990) showed that an attention-demanding task could modulate the duration of the aftereffect. Culham, Verstraten, Ashida, and Cavanagh (2000) showed that attentive tracking could modify the perceived direction of a directionally ambiguous stimulus—in their case a counterphasing radial grating. When tracking is maintained for some time in a given direction, a sudden cessation of tracking can result in a MAE opposite to the tracking direction, even though the physical motion signal is balanced across directions at all times. Such aftereffects are observed only when a dynamic test pattern is presented—no aftereffect is perceived on a static test pattern. So far, it is not known why that is the case.

STORAGE

The time that adapted neural substrates recover from adaptation can be delayed. Imagine that a MAE normally lasts *N* seconds. If an observer were to adapt to the same stimulus but close both eyes for *N* seconds after the adapting motion stopped, one would expect the MAE to have expired. However, in practice the MAE is present for several seconds after the eyes are reopened. This rather peculiar phenomenon is known as storage (Spigel, 1960). Storage depends on several parameters (see Thompson & Wright, 1994), and some even claim it is expected as an inherent part of the calibration processes of the visual system (van de Grind, van der Smagt, & Verstraten, 2004). New techniques (see next section) have shed light on the neurophysiological underpinnings of storage, but there is still no satisfactory explanation for this effect (Culham et al., 1999; He, Cohen, & Hu, 1998).

RECENT DEVELOPMENTS

The MAE remains the psychophysicist's microelectrode. However, with the advent of new neurophysiological techniques, in particular functional magnetic resonance imaging, the neurophysiological foundations of the MAE have become clearer. Tootell and colleagues (1995) found MAE-related activity in area MT+, an area that is responsive to motion. Huk, Ress, and Heeger (2001) showed that the story is somewhat more complicated as MT activity is modulated by attention. They engineered some experiments to disambiguate the effects of motion adaptation and attentional modulation of perceived direction.

It is unlikely that area MT is the only area responsible for MAEs. Several recent findings show that the MAE is not just a passive process but an active inference process based on recent sensory experience and current sensory input (Stocker & Simoncelli, 2008). For example, van der Smagt, Verstraten, and van de Grind (1999) showed that the same adapting motion stimulus could give rise to a number of different aftereffects, depending on the test pattern. Van der Smagt et al. employed random dot patterns in transparent motion, in which half the dots moved in a different direction and speed as compared to the other half—these stimuli appear as transparent surfaces that slide over one another. The direction of the aftereffect following adaptation to such transparent patterns depends on the nature of the dynamic noise test patterns, which appear like noise on a detuned TV set. When the adapting stimulus consists of a low- and a high-speed transparent motion pattern, the MAE is opposite the direction of the low-speed component when tested on a static pattern but opposite the direction of the high-speed pattern when tested with a high temporal frequency dynamic test pattern. Furthermore, the MAE can take one any direction between these extremes with adjustment of the temporal frequency of the test pattern.

The test pattern is also important when observers adapt to moving patterns defined by first-order or second-order statistics. First-order stimuli are defined by changes in luminance or chrominance, whereas second-order stimuli are defined by higher-order properties, such as contrast, depth, or relative motion, that are thought not to stimulate simple motion sensors directly. First-order adapting stimuli induce a MAE for both static and dynamic test patterns defined by first- or second-order properties. Second-order adapting stimuli induce a MAE only in second-order test patterns (Nishida & Sato, 1995). These data have been taken as evidence for independent adaptation at early (first-order) and late (first- and second-order) stages of visual processing.

ADAPTATION AND AFTEREFFECTS: A FUNCTIONAL ROLE?

Adaptation implies that an organism continuously adapts its sensory responses to changes in its ongoing environment. The simplest example is dark adaptation. When one turns off the light at night, the room appears pitch dark initially and nothing can be discriminated. However, after some time of adaptation, the few photons that are still

available to the visual system are enough to regain some vision in the dark. It looks as if the signal is amplified as the dynamic range or gain of photoreceptors adjusts to the prevailing light levels.

In this case, adaptation enables an organism to deal with a new or changing environment (see Webster, 2011). However, the functional role of the MAE is less clear. This visible aftereffect shows up after some time of adaptation, but adaptation effects can be measured after only milliseconds of adaptation (Kanai & Verstraten, 2005). Consequently, adaptation could well start as soon as there is a motion in the observer's field of view. For a moving observer, the retinal motion that is caused by self-motion—generating so-called optic flow—is important for navigation but is predictable based on the observer's own movement. What is more interesting for the visual system are the discontinuities in the flow field. In evolutionary terms, the source of discontinuities are often self-propelled organisms, which could mean food or danger. In either case, it is important that such objects capture attention and pop out. This process would be easier if the motion due to self-motion were discounted (Warren & Rushton, 2009), and it would be advantageous if this happened soon after there are changes in the movement of the observer. Future research will be required to understand the mechanisms and functional benefit of fast recalibrations to new environmental conditions.

REFERENCES

Addams, R. (1834). An account of a peculiar optical phenomenon seen after having looked at a moving body, etc. *The London and Edinburgh Philosophical Magazine and Journal of Science (Third Series)*, *5*, 373–374.

Aitken J. (1878). On a new variety of ocular spectrum. *Proceedings of the Royal Society of Edinburgh*, *10*, 40–44.

Anstis, S.M., & Moulden, B. P. (1970). After effect of seen movement: evidence for peripheral and central components. *Quarterly Journal of Experimental Psychology*, *22*, 222–229.

Anstis, S., & Reinhardt-Rutland, A. H. (1976). Interactions between motion aftereffects and induced movement. *Vision Research*, *16*, 1391–1394.

Anstis, S., Verstraten, F. A. J., & Mather, G. (1998). The motion aftereffect. *Trends in Cognitive Sciences*, *2*, 111–117.

Aristotle. (1955). *Parva naturalia* (W. D. Ross, ed.). Oxford: Clarendon Press.

Barlow, H. B., & Hill, R. M. (1963). Evidence for a physiological explanation for the waterfall phenomenon and figural aftereffects. *Nature*, *200*, 1345–1347.

Bex, P. J., Verstraten, F. A. J., & Mareschal, I. (1996). Temporal and spatial frequency tuning of the flicker motion aftereffect. *Vision Research*, *36*, 2721–2727.

Bowditch, H. P., & Hall, G. S. (1881). Optical illusions of motion. *Journal of Physiology*, *3*, 297–307.

Chaudhuri, A. (1990). Modulation of the motion aftereffect by selective attention. *Nature*, *344*, 60–62.

Culham, J. C., Dukelow, S. P., Vilis, T., Hassard, F. A., Gati, J. S., Menon, R. S., & Goodale, M. A. (1999). Recovery of fMRI activation in motion area MT following storage of the motion aftereffect. *Journal of Neurophysiology*, *81*, 388–393.

Culham, J., Verstraten, F. A. J., Ashida, H., & Cavanagh, P. (2000). Independent aftereffects of attention and motion. *Neuron*, *28*, 607–615.

Frisby, J. P. (1979). *Seeing: Illusion, brain and mind*. Oxford: Oxford University Press.

Gregory, R. L. (1981). *Mind in science: A history of explanations of psychology and physics*. London: Weidenfeld & Nicolson.

He, S., Cohen, E. R., & Hu, X. (1998). Close correlation between activity in brain area MT/V5 and the perception of a visual motion aftereffect. *Current Biology*, *8*(2226), 1215–1218.

Hess, R. F., Demanins, R., & Bex, P. J. (1997). A reduced motion aftereffect in strabismic amblyopia. *Vision Research*, *37*(10), 1303–1311.

Hiris, E., & Blake, R. (1992). Another perspective on the visual motion aftereffect. *Proceedings of the National Academy of Sciences USA*, *89*, 9025–9028.

Holland, H. C. (1965). *The spiral after-effect*. Oxford: Pergamon.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, *160*, 106–154.

Huk, A. C., Ress, D., & Heeger, D. J. (2001). Neuronal basis of the motion aftereffect reconsidered. *Neuron*, *32*, 161–172.

Kanai, R., & Verstraten, F. A. J. (2005). Perceptual manifestations of fast neural plasticity: Motion priming, rapid motion aftereffect and perceptual sensitization. *Vision Research*, *45*(25–26), 3109–3116.

Lucretius (1994). *De rerum natura/On the nature of the things* (R. E. Latham, trans.). London: Penguin. (Original work published 50 BC)

Mather, G. (1980). The movement aftereffect and a distribution-shift model for coding the direction of visual movement. *Perception*, *9*, 379–392.

Mather, G., Verstraten, F. A. J., & Anstis, S. M. (Eds.). (1998). *The motion aftereffect: A modern perspective*. Cambridge, MA: MIT Press.

Mollon, J. (1974, February). After-effects and the brain. *New Scientist*, 479–482.

Newsome, W. T., & Paré, E. B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *The Journal of Neuroscience*, *8*, 2201–2211.

Nishida, S., & Sato, T. (1995). Motion aftereffect with flickering test patterns reveals higher stages of motion processing, *Vision Research*, *35*, 477–490.

Over, R., Broerse, J., Crassini, B., & Lovegrove, W. (1973). Spatial determinants of the aftereffect of seen movement. *Vision Research*, *13*(9), 1681–1690.

Pantle, A. (1998). How do measures of motion aftereffect measure up? In G. Mather, F. A. J. Verstraten, & S. M. Anstis (Eds.), *The motion aftereffect: A modern perspective.* Cambridge, MA: MIT Press. 25–39.

Purkyně, J. E. (1820) Beiträge zur näheren Kenntniss des Schwindels aus heautognostischen Daten. *Medicinische Jahrbücher des kaiserlich-königlichen österreichischen Staates*, *6*, 79–125.

Riggs, L. A., & Day, R. H. (1980). Visual aftereffects derived from inspection of orthogonally moving patterns. *Science*, *208*(4442), 416–418.

Snowden, R. J., & Milne, A. B. (1996). The effects of adapting to complex motions: Position invariance and tuning to spiral motions. *Journal of Cognitive Neuroscience*, *8*(5), 435–452.

Spigel, I. M. (1960). The effects of differential post-exposure illumination on the decay of the movement after-effect. *Journal of Psychology*, *50*, 209–210.

Stocker, A., & Simoncelli, E. (2008). A Bayesian model of conditioned perception. *Advances in Neural Information Processing Systems*, *20*, 1409–1416.

Sutherland, N. S. (1961). Figural aftereffects and apparent size. *Quarterly Journal of Experimental Psychology*, *13*, 222–228.

Swanston, M. T., & Wade, N. J. (1992). Motion over the retina and the motion aftereffect. *Perception*, *21*, 569–582.

Thompson, P., & Wright, J. (1994). The role of intervening patterns in the storage of the movement aftereffect. *Perception*, *23*, 1233–1240.

Tootell, R. B. H., Reppas, J. B., Dale, A. M., Look, R. B., Sereno, M. I., Malach, R., Brady, T. J., Rosen, B. R. (1995). Visual motion aftereffect in human cortical area MT revealed by functional magnetic resonance imaging, *Nature*, *375*, 139–141.

van de Grind, W. A., van der Smagt, M. J., & Verstraten, F. A. J. (2004). Storage for free: A surprising property of a simple gain-control model of motion aftereffects. *Vision Research*, *44*, 2269–2284.

van der Smagt, M. J., Verstraten, F. A. J., & van de Grind, W. A. (1999). A new transparent motion aftereffect. *Nature Neuroscience*, *2*, 595–596.

Verstraten, F. A. J. (1996). On the ancient history of the direction of the motion aftereffect. *Perception*, *25*, 1177–1188.

Verstraten, F. A. J., Fredericksen, R. E., & van de Grind, W. A. (1994). The movement aftereffect of bi-vectorial transparent motion. *Vision Research*, *34*, 349–358.

Verstraten, F. A. J., Paffen, C. L. E., Carlson, T., van der Smagt, M. J., & Hogendoorn, H. (2007). Revealing recovery from adaptation below perceptual thresholds. *Perception*, *36*(Suppl.), 62.

Wade, N. J., & Verstraten F. A. J. (1998). Introduction and historical overview. In G. Mather, F. A. J. Verstraten, & S. M. Anstis (Eds.), *The motion-aftereffect: A modern perspective*: Cambridge, MA: MIT Press. 1–23.

Warren, P. A., & Rushton, S. K. (2009). Optic flow processing for the assessment of object movement during ego movement. *Current Biology*, *19*(18), 1555–1560.

Webster, M. A. (2011). Adaptation and visual coding. *Journal of Vision*, *11*(5):3, 1–23.

Weisstein, N., Maguire, W., & Berbaum, K. (1977). A phantom-motion aftereffect. *Science*, *198*(4320), 955–958.

Wohlgemuth, A. (1911). On the aftereffect of seen movement. *British Journal of Psychology Monographs*, *1*, 1–117.

Chapter 83
High Phi and Ghost Phi

Extreme Motion Illusions

Mark Wexler

HIGH PHI

We have recently discovered that when a moving stimulus is followed by certain visual transients, an illusion of very fast motion is perceived (Ito & Seno, 2009; Wexler, 2011; Wexler, Glennerster, Cavanagh, Ito, & Seno, 2013).[1] For example, consider the display in Interactive Figure IV.83-1 (the interactive video demonstrations can be seen at http://lpp.parisdescartes.cnrs.fr/wexler/compendium), in which a random texture undergoes a slow rotation. Every second most observers perceive a very fast jump in the direction opposite to the preceding or "inducing" rotation—as can be verified by reversing the inducing motion. These jumps are illusory: during the brief perceived jump, every frame is a new, random texture, uncorrelated with the previous textures; the last of these random textures is set to turning once again, and the sequence repeats. Thus there is no specific correspondence between the textures before, during, and after the jumps and no motion energy corresponding to the jumps. A general schema illustrating the displays is shown in Figure IV.83-1.

We call these illusory jumps "high phi." Observers generally perceive the jump as having a certain specific amplitude (typically between 45° and 90°; e.g., see Interactive Figure IV.83-1) and are sometimes surprised to learn that there is no correspondence between the new and old textures. The amplitude of the jump depends, among other factors, on the duration of the transient event, as shown in Interactive Figure IV.83-2: the longer the sequence of random textures, the farther back the illusory jump. However, when the transient lasts more than 100 to 200 ms, the jump becomes less salient. The perceived angular speed of the high phi jumps has been measured to be several hundreds of degrees per second (depending on the observer), and can easily exceed 10 times the speed of the inducing motion (Wexler et al., 2013).

Other visual transients can trigger the high phi jump. The illusion can be produced by suddenly changing local luminance, for instance by reversing contrast, as shown in Interactive Figure IV.83-3. Many observers can perceive high phi when simply blinking while watching motion. Projecting the inducing motion dichoptically—all in one eye and nothing in the other, or two halves of the pattern in different eyes—and suddenly switching the two eyes' images also produces the illusion. With a texture composed of discrete elements, suddenly changing the size or orientation of the elements leads to the high phi jump. Changing contrast by changing the luminance of the background does not produce the jump. Thus local monocular luminance changes are both necessary and sufficient to trigger the high phi illusion.

High phi can be either backward with respect to the preceding inducing motion or forward, depending on several factors. If the inducing motion is very brief, lasting less than 75 to 150 ms or so (the exact threshold varies from

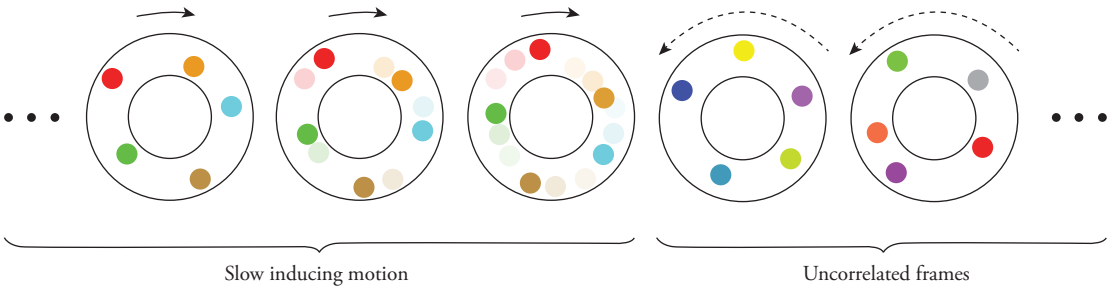
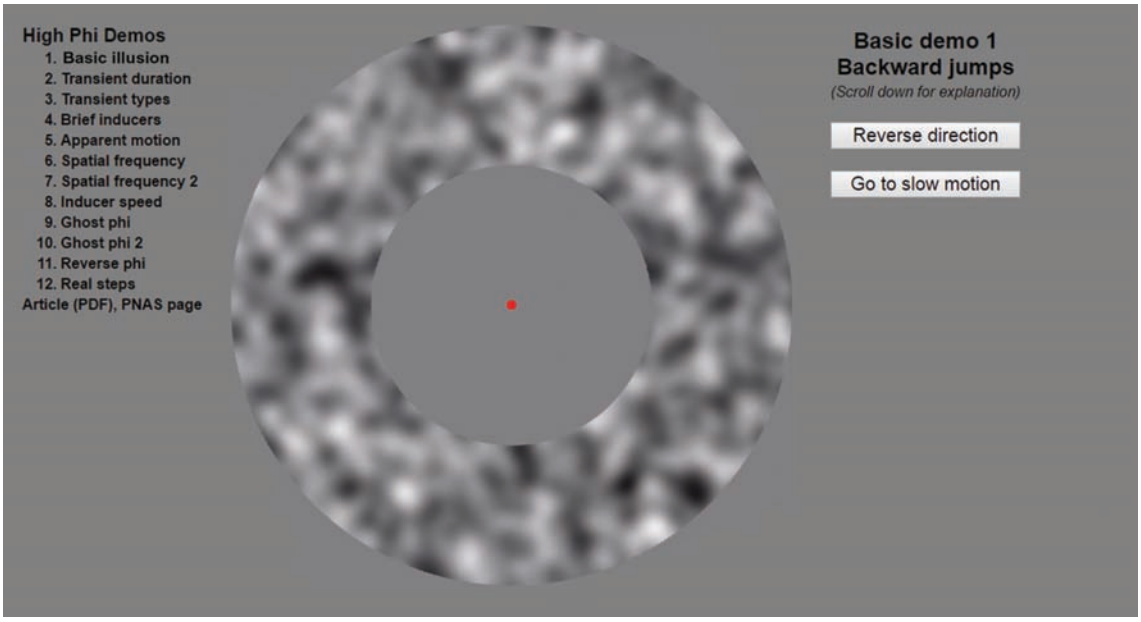


Figure IV.83-1. A schematic illustration of typical high phi displays (e.g., Interactive Figure IV.83-1 and IV.83-2). During the first phase, a texture is slowly rotated. Motion is perceived because each successive pair of frames is perfectly correlated, up to the small rotation. This is followed by a brief sequence of random, uncorrelated frames. This transient is perceived as a large, backward jump, provided the inducing motion lasts longer than 100 to 200 ms (as in Interactive Figure IV.83-1 and IV.83-2), or a large forward jump for briefer inducers (as in Interactive Figure IV.83-4).



Interactive Figures IV.83-1–IV.83-12. High phi and ghost phi demonstration.

observer to observer), the jump is forward, as can be seen in Interactive Figure IV.83-4 (Wexler et al., 2013). This demonstration shows that a tiny and very brief rotation, when immediately followed by a visual transient, is perceived as a much larger and faster rotation—but one that is always in the same direction as the tiny rotation (see the "slowed down" version in Interactive Figure IV.83-4 for an illustration of what actually occurs in the stimulus). A longer inducing rotation produces, as already shown, a jump in the opposite direction. The amplitudes of the forward and backward jumps vary considerably from one observer to another but are reported to be about equal to each other by each observer. Another way of controlling the direction of the jumps is through apparent motion: keeping the mean speed constant, brief steps lead to backward jumps and longer steps to forward jumps (see Interactive Figure IV.83-5). Finally, very low spatial frequencies, or viewing the stimulus in the visual periphery, systematically lead to forward jumps.

The amplitude of the high phi jump depends on the duration of the visual transient or noise, as mentioned and as shown in Interactive Figure IV.83-2. Another aspect of the stimulus that affects jump amplitude is spatial frequency. As shown in Interactive Figure IV.83-6, lower frequencies lead to larger jumps. In Interactive Figure IV.83-7, multiple spatial frequencies simultaneously present in the same stimulus show the effect of frequency in a different way: the inside of the ring, containing lower spatial frequencies, seems to jump faster and farther than the outside, leading to apparent deformation during the jumps (and to some degree during the real motion as well). The speed of the inducer, on the other hand, seems to have little effect on the illusion, as shown in Interactive Figure IV.83-8. For rotations, for example, there seems to be a wide range of inducer speeds, up to about 50°/s, ideal for producing the illusion. Within this range, the speed of the inducer has little or no effect on the perceived amplitude of the jumps. Very slow—indeed, even stationary—"inducers" can also produce the illusion. For instance, simply replacing one stationary random texture with another produces illusory jumps in some observers; however, these jumps are often described as being less salient and have arbitrary directions.

We have illustrated the high phi illusion using rotation, but similar effects can be obtained for other types of motion. Translation, shear, expansion, and contraction all produce illusory jumps in the direction opposite to the inducing motion. In all cases, the illusion is strongest when motion is several degrees from the center of the visual field. Multiple simultaneous high phi jumps, on the other hand, are difficult to obtain. Concentric counterrotating rings do not lead to high phi; disjoint rotating rings in different parts of the visual field yield less salient jumps.

GHOST PHI

If the transient is immediately "undone," restoring the texture back to its original state, a different kind of illusory motion is perceived. Consider, for instance, the stimuli in Interactive Figure IV.83-9. Here the inducing motion is followed by a transient (a polarity inversion or a new random texture, as before), but on the next frame, the transient is undone (by inverting the polarity once more, or by restoring the original texture). Although initially one may perceive a simple flash, many observers perceive an extremely fast, paradoxical motion opposite the direction of the inducer—paradoxical because it is impossible to say what, exactly, is moving. We call this illusory jump "ghost phi."

Compared to ghost phi, high phi (which can be seen in the same demonstration) looks more "concrete": one has the impression that something really has rotated. (Though, of course, nothing has rotated: as usual, there's no global correspondence between the two textures.) Most observers perceive ghost phi as even faster than high phi.

Another way of producing the ghost phi illusion is shown in Interactive Figure IV.83-10, based on the Shioiri-Cavanagh effect (Shioiri & Cavanagh, 1990). In the normal

version of the Shioiri-Cavanagh effect (which can be seen in the video), a texture is blanked for a brief period (here, two frames), then moved (here, a 5° rotation). Normally, a movement in the opposite direction is perceived, often with a larger amplitude—although the actual rotation can be seen by tracking a feature of the texture. The "ghost" version of this effect is identical, except that the rotation amplitude is smaller (here, 1°). In this version, many observers perceive a very fast and "ghostly" (i.e., nothing concrete seems to be turning) jump, always in the direction opposite of the 1° rotation.

RELATION TO MOTION ADAPTATION

The high phi illusion shares important features of motion adaptation, namely that long exposure to a motion stimulus leads to the motion aftereffect (MAE; see Anstis, Verstraten, & Mather, 1998), which is illusory motion in the direction opposite to the inducer, while brief exposure leads to motion priming (Pinkus & Pantle, 1997), where motion in the same direction as the inducer is favored. Moreover, the boundary between the motion priming and MAE regimes (about 200 ms; Kanai & Verstraten, 2005) is not far from the transition point that we find between forward and backward high phi (Wexler et al., 2013).

However, there are at least two important differences between high phi and the known forms of motion adaptation. There are two ways of measuring motion adaptation effects: using probes that are static or semi-static (e.g., moving perpendicular to the inducer) or using highly dynamic probes such as counterphasing gratings, whose motion is ambiguous. In the static MAE, one perceives very slow motion of the probe, slower than that of the adapter, and it is very difficult to report any sort of displacement because one does not perceive displacement, only motion. In the dynamic MAE (von Grünau, 1986), the perceived motion of the probe can be very fast, but this speed does not reflect any properties of the adaptation, only those of the probe: for a typical counterphased sine-wave grating probe, for example, the perceived speed is one-half the temporal frequency divided by the spatial frequency. Thus the speed of the dynamic aftereffect, though it can be fast, is imposed by the spatiotemporal properties of the probe.

In the case of high phi, the probe has a broad band of spatial frequencies and thus does not impose any particular speed. Nevertheless, most observers readily report perceived displacement, and these reports are consistent enough for the effect of various stimulus factors (transient duration, spatial frequency, etc.) to be easily measurable. The perceived speed of the high phi jump can be calculated as the slope of the linear regression of perceived displacement versus transient duration. For rotating stimuli with typical parameters, we find that the perceived angular speed of the high phi jump is several hundreds of degrees per second, with large variations between observers (Wexler et al., 2013).

Thus if high phi is related to motion adaptation, as it seems to be, then it shows that MAEs can be much faster than previously supposed (excluding cases where the speed of the aftereffect is determined by properties of the probe), with speed equal to several tens of times faster than that of the inducing motion. Aftereffects are known to increase as probe duration decreases (Wolfe, 1984), and our "probe"—the transient—is much briefer than the typical probes used to measure the MAE, but the speed difference is nonetheless striking. In addition, very little preceding motion is needed to obtain our illusion. In fact, a single tiny step (as little as 0.5 deg of rotation) on a single frame (typically 17 ms) preceding the transient is enough to produce to forward high phi jump, as can be seen in Interactive Figure IV.83-4 (Wexler et al., 2013). More recent work with a high refresh rate display shows that as little as 5 ms of preceding motion is sufficient.

RELATION TO REVERSE PHI

High phi seems to be related to the phenomenon called "reverse phi," discovered by Anstis (1970). A typical reverse phi stimulus can be seen in Interactive Figure IV.83-11, in which a stimulus is moved (here, rotated by 1°) and its contrast *simultaneously* inverted. A jump opposite to the small rotation is perceived. In a brief-inducer version of high phi, on the other hand, a stimulus is moved, and then its contrast is inverted on the *next* frame. A jump in the *same* direction as the rotation is perceived. Interactive Figure IV.83-11 allows one to compare these two closely related stimuli, both featuring a 1° rotation and a contrast inversion and differing only in whether these two transformations are simultaneous (reverse phi) or serial (high phi).

Many observers perceive the two jumps to be quite similar but of course in opposite directions. Both seem to be much larger than the 1° actual rotations, with the high phi jump often reported as somewhat larger. Because there are large variations across observers in the amplitude of high phi, we measured the amplitude of high phi and reverse phi jumps in a group of nine observers. Performing a linear regression of reverse phi versus high phi amplitudes, we found strong across-observer correlations between the two phenomena, with $R^2 \approx 0.9$.

The reverse phi phenomenon is often explained by reference to a pure sinusoidal grating, for which a phase change of by $+\theta$ together with a contrast inversion is identical to a phase change of $-(180° - \theta)$. Thus the phenomenon is trivial for a pure sine-wave grating but somewhat puzzling for a broadband stimulus such as the ones in Interactive Figure IV.83-11: each spatial frequency would seemingly yield a different displacement, with different directions. Why does reverse phi have the amplitude (and even the direction) that it does? In any case, the reverse phi and high phi phenomena seem to be related. Interestingly, a link between reverse phi and motion adaptation has been postulated on entirely different grounds (Bours, Kroes, & Lankheet, 2007).

RELATION TO GENERAL MOTION PERCEPTION MECHANISMS

Another way of producing the illusion is to superimpose actual jumps on the inducing motion. For example, consider 5° jumps, shown in Video IV.83-12. When the jumps

are either forward or backward with respect to the inducing motion, they are perceived veridically as such, namely small backward or forward jumps. However, when the amplitude of the jump is increased to, say, 20° (see Video IV.83-12), both backward *and* forward jumps are now perceived as backward! Moreover, if we increase the amplitude of the jump, all jumps continue to be perceived as backward, with roughly the same amplitude: the perceived amplitude of the 90° jump in Video IV.83-12 is not perceived as substantially larger than that of the 20° jump. These nonveridical backward jumps look the same as high phi jumps obtained by, say, randomizing the rotating texture, and that can also be seen in Video IV.83-12.

How can we make sense of this pattern of observations? Some experimental results can make them clearer. For jumps superimposed on a slow rotation, as in Interactive Figure IV.83-12, we asked subjects to report the direction and amplitude of the perceived jump. Results for one representative subject are shown in Figure IV.83-2. The subject perceives small jumps up to about ±12° veridically, showing a linear dependence of the perceived jump on the actual jump. (The slope is greater than 1, showing the subject's responses exaggerate the real jump by this factor, but the linear dependence shows that the responses reflect the actual jump, as does the sign on the response.) When the real jumps exceed about 12°, at which point they are perceived as being roughly 30° by the subject whose data are shown in Figure IV.83-2, there is a collapse. Nearly all jumps above +12° (forward) or below –12° (backward) are perceived as being backward, with an amplitude that no longer depends on the actual jump and that is roughly equal to –40°. Moreover, this latter amplitude is very close to that of the high phi jump, obtained by suddenly randomizing the texture (shown by the green dots on the right of Fig. IV.83-2). Thus the high phi jump is just outside of the range of displacements for which motion can be perceived veridically. In this sense, the high phi jump is not only fast, but it is the fastest motion observable!

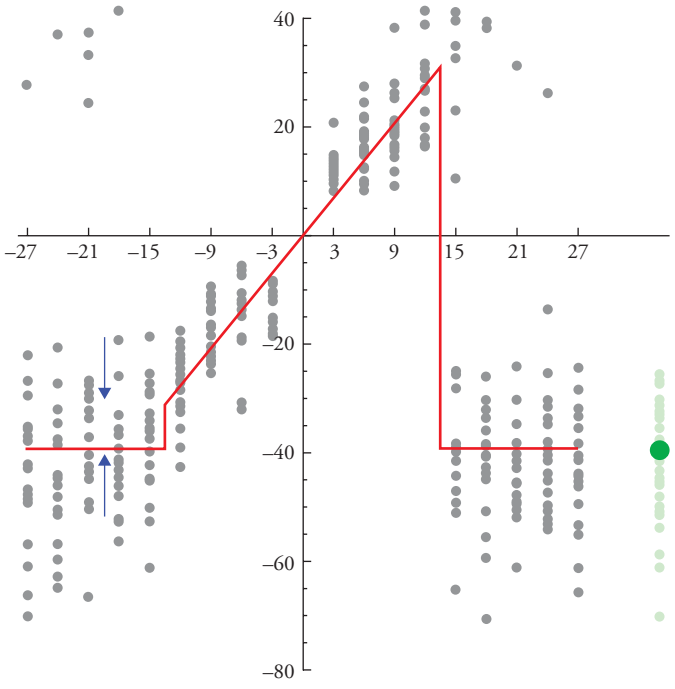


Figure IV.83-2. The- amplitude of the perceived jump as a function of the real jump, in degrees. The jumps were superimposed on a slow rotation. Positive angles are in the same direction as the rotation; negative angles are opposite. The gray dots denote individual trials and the red curve a piecewise-linear fit. Pale green dots on the right show trials on which, instead of jump, the texture was randomized—the standard high phi effect; the large green dot shows the mean perceived jump on these trials. The blue arrows show the distance between the largest veridically perceived jump and the default backward jump perceived at larger amplitudes.

These observations suggest the following possible explanation for the illusion. Motion perception operates through what is effectively a search for local correspondence, limited in space by an extent usually referred to as d_{max} (Braddick, 1974). When the transient occurs, the search begins, but no correspondence is found, either because the post- and pre-transient textures are never positively correlated in our stimuli or because they are correlated with a rotation above d_{max} (for the large-step stimuli described in this section). Thus, with a certain range of motions having been *excluded*, the visual system then applies the well-known minimal motion principle, namely that perceived motion is the slowest possible motion compatible with sensory evidence (Ullman, 1979). Here this perceived motion would be just outside the searched range and therefore above d_{max}. This idea predicts that the amplitude of the high phi jump should lie just outside of d_{max}, as it does in the data shown in Figure IV.83-2. It does not predict the direction of the high phi jump, though the effects of motion adaptation on this process are plausible.

This explanation for high phi is supported by its dependence on spatial frequency, namely the increase of perceived illusory motion as spatial frequency decreases, as discussed previously and seen in Interactive Figures IV.83-6 and IV.83-7. The spatial limit on motion perception, d_{max}, depends on spatial frequency in much the same way (Chang & Julesz, 1983). However, there are some difficulties with the explanation as well. For instance, why does high phi motion continue as the duration of the transient increases? This continuation seems to imply the existence of a recurrent or feedback process.

More generally, the transient noise in high phi can be seen as an amplifier, boosting tiny motion signals to their saturation values. For the brief-inducer version the amplified signal is the motion itself; for the long-inducer version it is the activity due to adaptation. The existence of an amplifier in the motion-perception system suggests the presence of positive feedback loops, which have been proposed (Neri, 2014) on totally different grounds.

ACKNOWLEDGMENTS

We wish to thank Patrick Cavanagh, Thérèse Collins, Andrew Glennerster, Hiroyuki Ito, and Frans Verstraten for useful conversations and Marianne Duyck and Hanako Ikeda for help in data collection.

NOTE

1. The high phi illusion presented here was independently discovered by Ito and Seno (2009) and Wexler (2011). To our knowledge, the only similar effect ever published was the largely forgotten study of Green, Chilcoat, and Stromeyer (1983).

REFERENCES

Anstis, S. (1970). Phi movement as a subtraction process. *Vision Research*, *10*, 1411–1430.

Anstis, S., Verstraten, F. A. J., & Mather, G. (1998). The motion aftereffect. *Trends in Cognitive Sciences*, *2*, 111–117.

Bours, R. J., Kroes, M. C., & Lankheet, M. J. (2007). The parallel between reverse-phi and motion aftereffects. *Journal of Vision*, *7*(11):8, 1–10.

Braddick, O. (1974). A short-range process in apparent motion. *Vision Research*, *14*(7), 519–527.

Chang, J. J., & Julesz, B. (1983). Displacement limits for spatial frequency filtered random-dot cinematograms in apparent motion. *Vision Research*, *23*(12), 1379–1385.

Green, M., Chilcoat, M., & Stromeyer, C. F.III. (1983). Rapid motion aftereffect seen within uniform flickering test fields. *Nature*, *304*(5921), 61–62.

Ito, H., & Seno, T. (2009). Reversed motion perception during blinks. *Perception*, *38*(ECVP Abstract Suppl.), 90.

Kanai, R., & Verstraten, F. A. J. (2005). Perceptual manifestations of fast neural plasticity: Motion priming, rapid motion aftereffect and perceptual sensitization. *Vision Research*, *45*, 3109–3116.

Neri, P. (2014). Dynamic engagement of human motion detectors across space-time coordinates. *Journal of Neuroscience*, *34*(25), 8449–8461.

Pinkus, A., & Pantle, A. (1997). Probing visual motion signals with a priming paradigm. *Vision Research*, *37*, 541–552.

Shioiri, S., & Cavanagh, P. (1990). ISI produces reverse apparent motion. *Vision Research*, *30*(5), 757–768.

Ullman, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.

Wexler, M. (2011). An ultra-fast motion aftereffect. *Journal of Vision*, *11*(11), 756.

Wexler, M., Glennerster, A., Cavanagh, P., Ito, H., & Seno, T. (2013). Default perception of high-speed motion. *Proceedings of the National Academy of Sciences USA*, *110*, 7080–7085.

Wolfe, J. M. (1984). Short test flashes produce large tilt aftereffects. *Vision Research*, *24*(12), 1959–1964.

Chapter 84
Stereokinetic Phenomena

Mario Zanforlin

Stereokinetic phenomena first observed by V. Benussi and investigated by C. Musatti (1924, 1955) are produced by two-dimensional rigid figures drawn or stuck on a platform in slow rotatory motion (10 to 20 cycles per minute) in the frontal plane. After a certain period of observation, the figures produce a much stronger three-dimensional impression than that produced by perspective, as they emerge from the flat platform and appear as solid and real. The percepts are largely independent of the rotational speed of the stimuli, of their location on the rotating platform, and of their distance from the observer.

A circle with an eccentric dot inside it appears as a cone of well-defined height. An ellipse of uniform color appears first as a deforming figure, then as a rigid disc slanted in depth (Musatti, 1924, 1955), and then as a solid ellipsoid with its major axis slanted in depth (Mefferd, 1968; Vallortigara, Bressan, & Zanforlin, 1986; Zanforlin, 2000). All the various appearances of the ellipse alternate with prolonged observation (Vallortigara, Bressan, & Bertamini, 1988). A rotating bar of constant length appears slanted in depth and 1.5 times longer than when it is stationary on the frontal plain (Zanforlin & Vallortigara, 1988). A circle and a semicircle appear to connect, forming an anomalous solid cylinder (Zanforlin & Vallortigara, 1990). All these moving figures can be produced on a computer screen, but the effect is not as striking, because on the computer everything appears behind the screen and the impression of a "real object" is lost. Original stereokinetic phenomena are so beautiful that they have been presented as works of art such as that by M. Duchamp (1935) in *Rotorelief*.

Three aspects of these objects are theoretically interesting: (a) they do not appear to rotate but describe a circular translation (a movement analogous to that of a hand drawing a circle with the thumb oriented to the left and all its points moving at the same speed); (b) they appear to be three-dimensional and solid; and (c) they appear of a well-defined length in depth. The first two aspects were explained by Musatti (1924, 1955) and others as due to different perceptual mechanisms. In the first case, the apparent translation has been considered as a general "tendency to constant orientation in space" of all the percepts (Musatti did not quote any evidence) or to limited aperture of peripheral movement receptors due to the curved contour ("misperception"; Ullman, 1984b). In the second case, it has been considered as "constancy of form" (Musatti, 1924), similar to the "rigidity assumption" (Ullman, 1979, 1984a). The third aspect of these objects (i.e., their apparent length in depth) cannot be explained by either of these hypotheses. However, Zanforlin (1988, 2000) has demonstrated that all three aspects are connected and can all be explained by a hypothetical perceptual process that minimizes both apparent speeds and relative differences between the speeds of all the points of the rotating configuration. For example, two different speeds on the frontal plain can be equated or the difference minimized by adding a depth component to one of them. A more formal demonstration can be found in Beghi, Xausa, and Zanforlin (2008).

The most interesting of the phenomena from a theoretical point of view is the rotating ellipse, which has been analyzed by various authors in addition to Musatti, including Mefferd (1968), Hildreth (1984), Todorović (1993), and Weiss and Adelson (2000).

An ellipse of average eccentricity (i.e., not too similar to a circle or a line) and uniform surface when set in rotation on the frontal plane appears at first just as it is: a rotating rigid and flat ellipse (Zanforlin, 2000). After few seconds, it appears to deform while the contour points maintain their orientation in space; that is, the points at the top and the bottom, on the left and on the right, maintain their position during rotation of the axis, as they can easily be perceived in spite of the uniform contour. Wallach (1990) speaks of "identity imposition." This effect, as pointed out already, has been explained as a general tendency of all percepts (Musatti, 1924). As the apparent orientation of the contour points can be seen as a counterrotation in relation to the rotating axis of the ellipse, the apparent orientation can be considered the result of a minimization process that equalizes their average speeds to that of the center of the ellipse.

As a result of their orientation, the contour points find themselves corresponding alternately to the minor and major rotating axes, appearing to approach and recede from the center: the ellipse appears as deforming or elastic (see Fig. IV.84-1).

Continuing observation after deformation, we find the ellipse appears as a rigid disc tilting in depth. This second appearance is particularly interesting because it can be explained in various ways: by an equalization of the distances of the contour points from the center of the ellipse (the

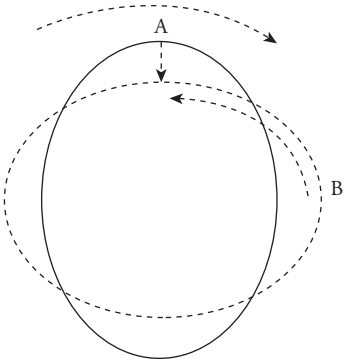


Figure IV.84-1. The contour points of a rotating ellipse appear to maintain their orientation in space, and the ellipse appears as elastic or as a rigid disc. The dotted arc above the ellipse indicates the direction of rotation. A is the point corresponding to the extremity of the longer axis, and it should move to B during a quarter rotation. But due to the uniformity of the contour, it can make a counterrotation, indicated by the dotted arc inside the ellipse, and appears to maintain its position at the top, moving up and down in correspondence with the rotating axis. (Adapted from Musatti, 1924.)

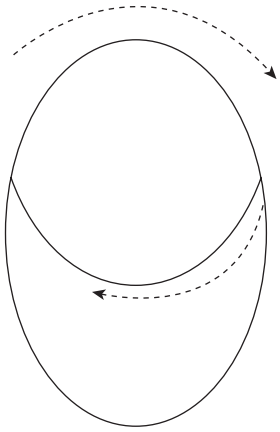


Figure IV.84-2. An ellipse with a semicircle. When set in rotation, the surface of the ellipse appears to maintain its orientation in space by counterrotating around the major axis. The ellipse appears as a solid ellipsoid, slanted in depth and longer than when it is stationary on the frontal plane. (Adapted from Zanforlin, 2000.)

rigidity assumption or tendency to constancy; Hullman, 1984a; Musatti, 1924), by a vector analysis (Todorović, 1993), and as a further minimization of the speed differences of contour points (i.e., speed becomes uniform) around the center by addition of a depth component (Beghi et al., 2008; Zanforlin, 2000; see Fig. IV.84-1).

The third appearance is as a solid ellipsoid (like an egg or a rugby ball), and if a semicircular segment is added around the major axis of the ellipse, this will be its unique appearance (see Fig. IV.84-2).

This appearance cannot be explained by a rigidity assumption, which implies a minimization or an equalization of distance differences of the contour points from the center, because the major axis of the ellipse appears longer in depth than when it is stationary on the frontal plane.

A possible explanation is to suppose that the visual system equalizes or minimizes the speed differences among the contour points. This hypothesis derives from the Gestalt general principle of minimum differences in perception. Musatti (1931) has demonstrated that all of Wertheimer's laws are instances of this minimum principle.

If we suppose that the contour points move across the uniform surface, as indicated by the semicircle in Figure IV.84-2, they make a circular counterrotation around a correspondent point of the major axis, completing their trajectory amodally. In this way they appear to maintain their orientation in space and equalize their speed to that of the correspondent point on the major axis. Obviously the points of the major axis have different speeds depending on their distance from the rotation center. However, equalizing the different speeds of the ellipse axis is equivalent to equalizing the speeds of all the points of a rotating line segment or a bar of constant length.

A bar of constant length that rotates on the frontal plane at first appears to make strange movements, but after continued observation it appears as a bar slanted in depth, longer than when it is stationary, and describing a cone surface. As the length of a bar on the frontal plane is constant, it is, by definition, a rigid "perceptual object," and its apparent displacement in depth cannot be explained by a rigidity assumption. However, this displacement can be explained by equalizing the different speeds of the points of the bar, as illustrated schematically in Figure IV.84-3 (Beghi et al., 2008; Zanforlin, 2000).

Before entering into a possible explanation, two considerations should be kept in mind. First, the rotating bar on the frontal plane can be the two-dimensional projection of a bar of any length tilted in depth. Thus two percepts are possible: a bar lying on the frontal plane or a bar slanted in depth. Second, two points that move at different speeds on the frontal plane can equalize their speed by the addition of a depth component to the slower one. The addition of a depth component will displace the point in depth when the movement stops.

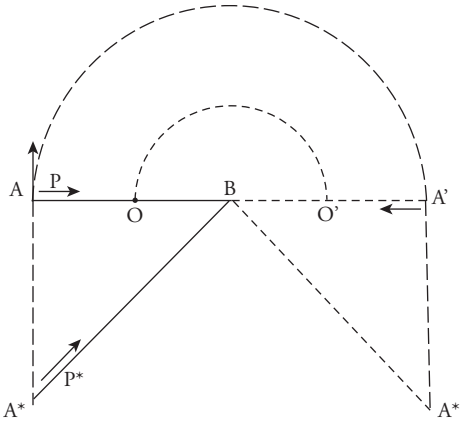


Figure IV.84-3. Two-dimensional schematic representation of a rotating bar of constant length moving around one of its extremities. A hypothetical point P moving from A to B in the first half rotation and from B to A in the second half (the latter movement is not represented). The bar can be considered the two-dimensional projection of a longer bar slanted in depth drawn in the lower half of the figure as seen from above. (Adapted from Zanforlin, 2000.)

We can use different mathematical procedures to equalize the different speeds of the points of a bar rotating around one of its extremes (at the moment it is not known how the visual system achieves this equalization). But is possible to demonstrate that the tilted bar percept is the result of an equalization process.

One possibility is to consider a "virtual point" P (a mathematical device not phenomenally visible) that moves from A to B along the bar on the frontal plain during the first half rotation and back from B to A in the second half rotation. During this movement the speed of point P will correspond to the different speeds of the various points of the bar. But the speed of the central point is special, as it can be considered the "common speed" of all the points of the bar (Cutting & Proffitt,1982; Johansson, 1974). When the bar is stuck in any other position on the rotating platform, it will appear as rotating around its center with the extremities acquiring equal speed around the center.

It is possible to equalize the speed of point P and that of all the points of the rotating bar to that of point O. As the length of the trajectory of point O in half rotation is $1/2\pi$, the distance that the virtual point P moving at the same uniform speed will also be $1/2\ \pi$. No matter that the trajectory of O is semicircular while that of P is rectilinear as it moves along the bar; at an equal speed for an equal time, the distance traveled will be equal.

As the bar on the frontal plane is of length 1, point P cannot describe a trajectory of length 1.5 on the frontal plane; it must travel along the bar perceived as slanted in depth from A* to B*. All the points of the bar can move at the same speed as P up and down the tilted bar (like the carriages of a train), describing a trajectory of the same length, turning when they reach the bar extremities. Now if all the points are stopped (we can eliminate the virtual point P), they will maintain their relative depth position along the tilted bar and the bar will appear of the same length and slanted in depth. As pointed out previously, the different speeds of the points of the bar perceived as lying on the frontal plane are equalized by a proportional depth displacement effect (Zanforlin, 2000).

An alternative way to equalize the different velocities of the points of the rotating bar is to start by adding a depth component to each of them, a calculation that leads to approximately the same result. For a detailed analytical demonstration, see Beghi et al. (2008).The experimental results are in good agreement with theoretical predictions, see also Zanforlin and Vallortigara (1988).

The results of this speed difference minimization or equalization is that the rotation on the frontal plane of the bar and of the ellipsoid is transformed into a translatory movement in the third dimension, which can be called a circular oscillation or a circular tilting—a movement biologically more compatible than a rotation. This movement is analogous to that of the hand–elbow joints when tracing a circle on the blackboard, when the thumb of the right hand points continuously to the left. The direction of the thumb, as mentioned before, demonstrates that the hand is making a circular translation. Neither the hand nor the arm can rotate around the elbow to draw a full circle. The elbow can bend or oscillate up and down, left and right, and describe a circle.

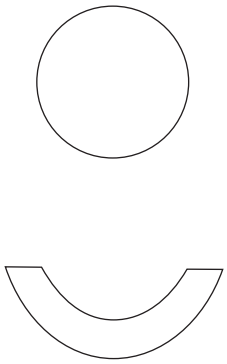


Figure IV.84-4. A circle above a semicircle that forms an anomalous cylinder when set in rotation. (Adapted from Zanforlin & Vallortigara, 1988.)

The same explanation applies to the cylinder formed by two circles that rotate around each other. They appear displaced in depth as the two extremities of an invisible thick bar or cylinder. However, the invisible bar or cylinder can be made visible as anomalous contours by deleting half of one of the circles (see Fig. IV.84-4).

There are various other phenomena all based on circles, ellipses, or bars, rotating on the frontal plain that produce three-dimensional percepts that can be explained by a minimization of velocity differences. This hypothesis does not contradict the minimization of distance differences derived from the tendency to constancy or the rigidity assumption. These two hypotheses appear complementary, as both can be derived from the general minimum principle of Gestalt theory.

REFERENCES

Beghi, L., Xausa, E., & Zanforlin, M. (2008). Modelling stereokinetic phenomena by a minimum relative motion assumption: The tilted disk, the ellipsoid and the tilted bar. *Biological Cybernetics*, *90*, 115–123.

Bressan, P., & Vallortigara, G. (1986). Multiple 3-D interpretations in a classic stereokinetic effect. *Perception*, *15*, 405–408.

Cutting, J. E., & Proffitt, D. R. (1982). The minimum principle and the perception of absolute, common and relative motion. *Cognitive Psychology*, *14*, 211–246.

Duchamp, M. (1935). *Rotorelief*. New York, Guggenheim Library.

Hildreth, E. C. (1984). Computation underlying the measurement of visual motion. *Artificial Intelligence*, *23*, 309–354.

Johansson, G. (1974). Vector analysis in visual perception of rolling motion: A quantitative approach. *Psychologische Forschung*, *36*, 311–319.

Musatti, C. L. (1924). Sui fenomeni stereocinetici. *Archivio Italiano di Psicologia*, *3*, 105–120.

Musatti, C. L. (1931). Forma e assimilazione. *Archivio Italiano di Psicologia*, *9*, 61–156.

Musatti, C. L. (1955). La stereocinesi e il problema della struttura dello spazio visibile. *Rivista di Psicologia*, *49*, 3–57.

Mefferd, R. B. Jr. (1968). Perception of depth in rotating objects: 7. Influence of attributes of depth in stereokinetic percepts. *Perceptual and Motor Skills*, *27*, 1179–1193.

Todorović, D. (1993). Analysis of two- and three-dimensional rigid and non-rigid motions in the stereokinetic effect. *Journal of the Optical Society of America A*, *10*(5), 804–826.

Ullman, S. (1979).*The interpretation of visual motion*. Cambridge, MA: MIT Press.

Ullman, S. (1984a). Maximizing rigidity: The incremental recovery of 3-D structure from rigid and non-rigid motion. *Perception*, *13*, 255–274.

Ullman, S. (1984b). Rigidity and misperceived motion. *Perception*, *13*, 218–219.

Wallach, H., & Centrella, N. M. (1990). Identity imposition and its role in a stereokinetic effect. *Perception & Psychophysics*, *48*, 535–542.

Weiss, Y., & Adelson, E. H. (2000). Adventure with gelatinus ellipses—Constraints on models of human motion analysis. *Perception*, *29*, 543–566.

Vallortigara, G., Bressan, P., & Bertamini, M. (1988). Perceptual alternation in stereokinesis. *Perception*, *17*, 31–34.

Vallortigara, G., Bressan, P., & Zanforlin, M. (1986). The "Saturn" illusion: A new stereokinetic effect. *Vision Research*, *26*, 811–813.

Zanforlin, M. (1988). The height of a stereokinetic cone: A quantitative determination of a 3D effect from 2D moving patterns without a "rigidity assumption." *Psychological Research*, *50*, 162–172.

Zanforlin, M. (2000). The various appearances of a rotating ellipse and the minimum principle: Review and an experimental test with non-ambiguous percepts. *Gestalt Theory*, *22*, 157–184.

Zanforlin, M., & Vallortigara, G. (1988). Depth effect from a rotating line of constant length. *Perception & Psychophysics*, *44*, 493–499.

Zanforlin, M., & Vallortigara, G. (1990). The magic wand: A new stereokinetic anomalous surface. *Perception*, *19*, 447–457.

Chapter 85
Motion Illusions in Static Patterns

Johannes M. Zanker

INTRODUCTION

Some op art paintings, such as Bridget Riley's *Fall* (1963, Tate Gallery London), create vivid sensations of shimmering and movement, despite the fact that they are nothing more than simple static patterns of paint on a static canvas. Similar examples of high-contrast periodic patterns (*gratings*), which are arranged such that they do not form straight lines (*phase modulation*), have been most effectively and systematically employed by Riley but have also been studied by other modern and contemporary artists, including Vasarely, Klee, Ouchy, and Leviant. Perhaps more surprisingly, they can also be found in a much wider range of art forms, including traditional weaving patterns and aboriginal bark paintings (Fig. IV.85-1a).

What could be the cause for such dynamic illusions? A number of visual scientists (e.g., MacKay, 1957; Wade, 1982), fascinated and puzzled by such motion sensations in static images, went to the drawing board to design their own versions of motion illusions (Fig. IV.85-1b), somewhat resembling a quiet counterpoint to a busy world of physically moving images and dynamic animations. All such stimuli are generated from spatial displacements of image elements that happen over time, which through their inherent spatiotemporal correlation create *motion energy*—which is regarded as the essence of perceived motion. For a couple of decades, scientists have been speculating whether the instability of the eye, endlessly shifting the gaze across the static patterns and changing focus, could be responsible for the perceived motion (Campbell & Robson, 1958; Gregory, 1993; MacKay, 1957) or whether higher cortical mechanisms are required to explain the phenomenon (Zeki, Watson, & Frackowiak, 1993). By combining systematic perceptual experiments, observations of eye movements, and computer simulations of the brain activity under such conditions, recent work (Troncoso, Macknik, Otero-Millan, & Martinez-Conde, 2008; Williams, Zanker, & Ashida, 2005; Zanker, Doyle, & Walker, 2003) has led to

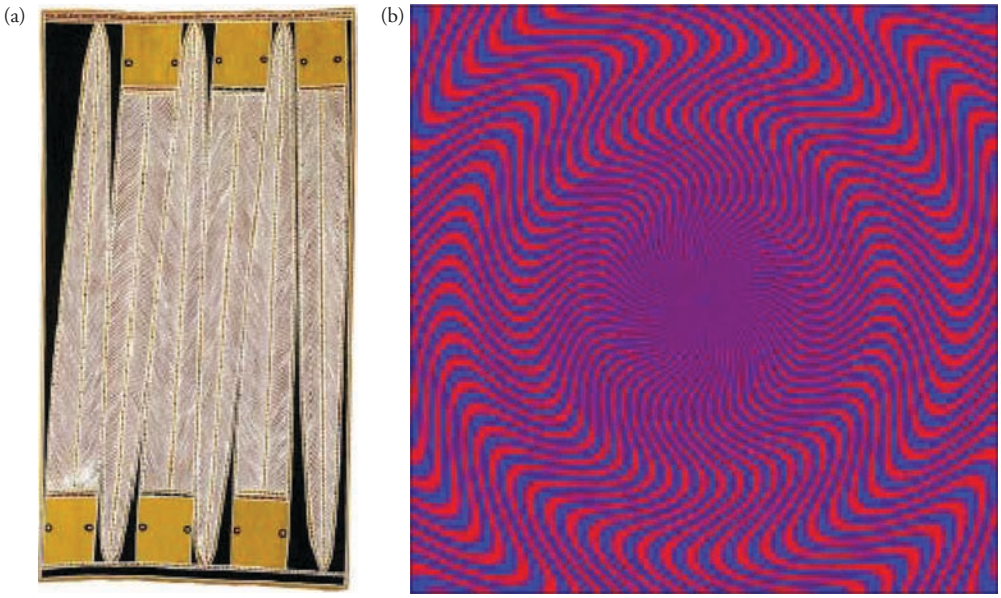


Figure IV.85-1. Motion illusions in (a) Aboriginal arts from Ramingining, Australia (Jimmy Wululu *Giny'giny*, c. 1984) (Figure by which is taken from the Museum of contamorary Art, Sydney, catalogue: The Native Born, Authors: Djon Mundine, John Rudder (Pub: Museum of Contemporary Art- Australia, 2000, Pbk, 242 pp) and (b) stimulus patterns generated by scientist for psychophysics experiments (riloids, JMZ), inspired by op art. (Figures by Johannes M. Zanker.)

the conclusion that such op art effects rely on involuntary shifts of the retinal image that are picked up by low-level motion detectors. This line of thought is described in more detail in the first part of this chapter.

The second part of the chapter then traces the similarities and differences between motion illusions generated by phase-modulated gratings with a second group of illusions that involve regular patches that are filled with oriented textures characterized by shades of grey or color, also known as *spatial gradients*. Arguably the most famous of these stunning sensations is the rotating snake illusion designed by Akiyoshy Kitaoka (2002) which went viral on the Internet (http://www.ritsumei.ac.jp/~akitaoka/index-e.html) at the turn of this century. This stimulus shares some similarities with concentric arrangements of some phase-modulated gratings (see Fig. IV.85-1b), and in an attempt to exactly understand the information-processing mechanisms underlying this phenomenon, it has been "simplified" in a few different ways (see, e.g., Fig. IV.85-3), suggesting a distinct role for dynamic background changes, or *temporal modulation*. The ultimate goal for vision science should be to arrive at a plausible computational model that embodies such mechanisms and allows us to account for existing illusions and generate novel ones and in doing so help us to understand how the brain works.

PHASE-MODULATED GRATINGS AND EYE MOVEMENTS

The scientific study of motion illusions in static patterns starts with the careful construction of stimuli, which can be manipulated to assess under which conditions the illusion is observed and when it disappears. In an attempt to start from a pattern that is composed from simple elements that can be defined by a small set of analytical expressions, Bridget Riley's *Fall* was crudely captured by phase-modulated, grey-level sine-wave gratings with three parameters (Zanker, 2004) that opens up a large (and easy to quantify) perceptual space for such *riloid* patterns, which resemble Riley's painting to a variable degree and can be extended in different geometrical and color configurations such as shown in Figure IV.85-1b.

The second step for the visual scientist is to evaluate the existence, quality, and strength of the illusion, ideally providing a quantitative measure of the perceptual response that can be related to the stimulus properties. The simplest form of such measurement is collecting subjective ratings of observers reporting the strength of the perceived illusion on a Likert scale. While such judgements allow a quick assessment of strong effects, and therefore have been used to establish the basic conditions under which the motion illusion is perceived (Zanker & Walker, 2004), it is clear that they lack the accuracy and reliability of exact quantitative measurements. More sophisticated comparison methods, involving two-alternative forced-choice decision tasks and other psychophysical paradigms, have been developed and used to confirm the initial observations and characterize the optimal stimulus parameters to generate the illusion (Zanker, Hermens, & Walker, 2010).

Extended and lively debates about the origin of motion illusions in a range of similar configurations of periodic line patterns were inspired by suggestions of Moirés-like shimmer experienced in radial line patterns (MacKay, 1957) and by more subtle motion interactions between regions covered with different textures (Leviant, 1996). This debate culminated in the question of whether more peripheral causes, like movement or accommodation of the eye (e.g., Campbell & Robson, 1958; Gregory, 1993), or more central effects, based on cortical mechanisms (e.g., Zeki, 1994), are crucial for the illusion. Recent improvements in eye-tracking technology facilitated a number of studies that investigated the relationship between this motion illusion and eye movements (Zanker & Walker, 2004), in particular small, involuntary saccades (*micro-saccades*) that are made frequently during periods of fixation (Martinez-Conde & Macknik, 2007). In such experiments, the timing of illusory percepts could be linked to the appearance of micro-saccades (Troncoso et al., 2008) and, in the case of riloids, also provided an opportunity to quantitatively relate the strength of the perceived illusion to the expected motion signals resulting from saccadic displacements (Zanker et al., 2010).

To fully understand the role of retinal image displacements for generating these illusions, a biologically plausible computational model of the cortical mechanisms of motion information processing was used to derive quantitative predictions of the motion energy created by different variants of gratings (Zanker, 2004). Whereas it is intuitively clear that displacements of phase-modulated gratings can lead to apparent motion signals that point in different directions (as indicated by the green and red arrows in the top of Fig. IV.85-2b, the closest match between lines before and after a horizontal displacement can point to the right or to the left), it requires a two-dimensional motion detector network (2DMD model: a two-dimensional array of local motion detectors that indicate the direction and strength of motion responses at any given location) to appreciate the full variation of local motion signals generated by micro-saccades for such stimuli: the wide range and the stripe pattern of local motion directions for a random micro-saccadic displacement (shown as different colors in the lower part of Fig. IV.85-2a) resemble the incoherent variation of strong local directions experienced when viewing a riloid (top of Fig. IV.85-2a).

Control condition with a diagonal displacement of a radial line pattern (top), leading to a motion signal map (bottom) with coherent direction gradients in the peripheral regions and inversions/distortions in the central region of the stimulus.

Joining together the psychophysical, oculomotor, and modeling results leads to a simple proposition that explains the motion illusion with minimal assumptions: involuntary micro-saccades in combination with standard cortical motion processing mechanisms generate characteristic motion signal distributions that correspond to the perceived illusion. Following this line of thought, however, one might wonder why for other stimuli such as simple checkerboard patterns, straight gratings, or the outer regions of the radial gratings shown in Figure IV.85-2c, no such illusory motion is perceived. The 2DMD response maps shown in Figure IV.85-2c can provide a clue to resolve this conundrum: in contrast to the incoherent response maps in Figure

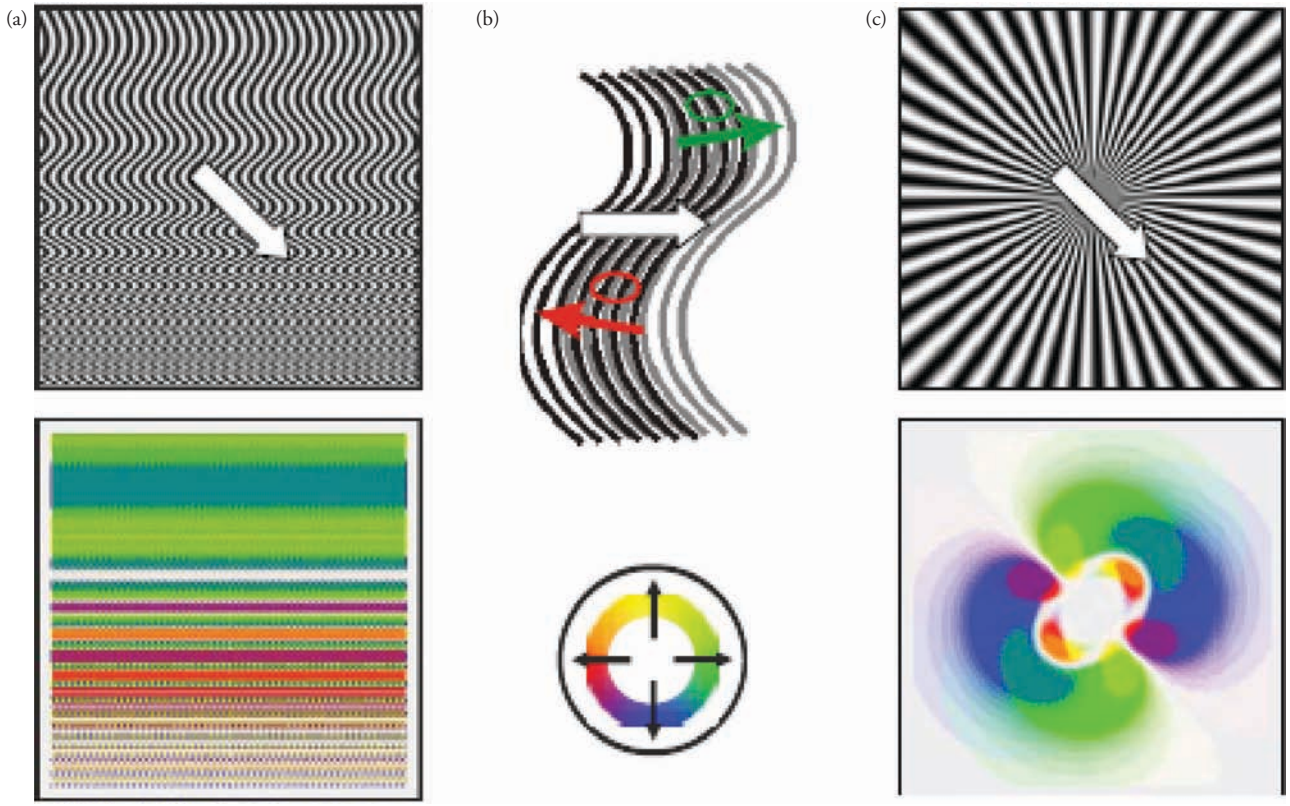


Figure IV.85-2. A functional explanation of motion illusions. (a) Riloid stimulus (top) and distribution of local motion signal (bottom, in color code; see bottom inset of Fig. IV.85-2b) detected for a diagonal micro-saccade (see white arrow). (b) Matching patterns of closest neighborhood for a lateral saccadic displacement of sine-wave lines (top, white arrow) and schematics color circle to indicate the local motion direction detected by the motion detector network (2DMD model) is encoded in motion signal maps (bottom). (c) Control condition with a diagonal displacement of a radial line pattern (top), leading to a motion signal map (bottom) with coherent direction gradients in the peripheral regions and inversions/distortions in the central region of the stimulus. (Figures by Johannes M. Zanker.)

IV.85-2a, the outer regions of the radial line show a systematic variation of motion signals that on average provide a reliable neural signal of the overall pattern displacement. Such signals can be used by the brain to correct for the overall movement of the eye, using the cancellation mechanisms that are well known for larger saccades. The use of coherent motion signals that indicate the overall stimulus displacement resulting from micro-saccades can explain perceptual stability of a scene under normal conditions, which breaks down for the incoherent motion signal distributions arising from phase-modulated gratings leading to the observed illusion of disjointed local pattern motion.

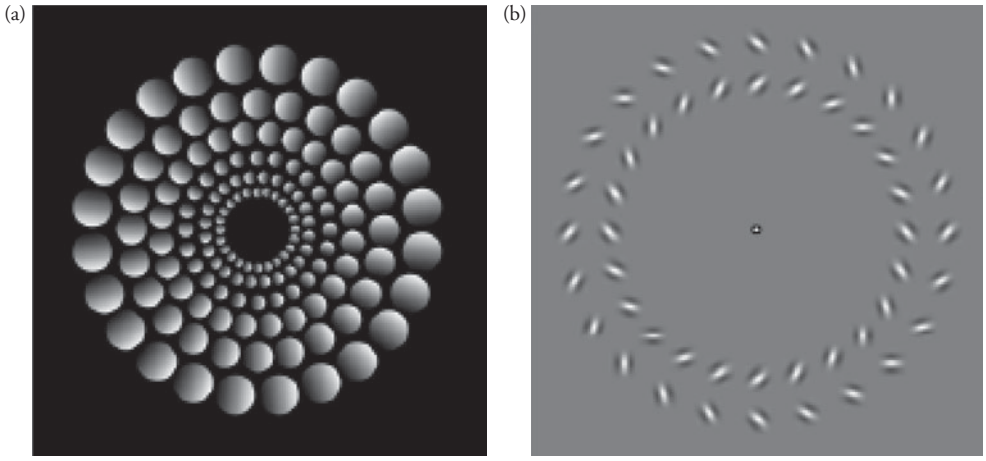


Figure IV.85-3. Static motion illusions from local brightness gradients. (a) The spinning disk illusion, a simplification of a rotating snake in black and white. (b) One version of the Pinna-Brelstaff illusion, with two circles of tilted Gabor patterns that begin to rotate when the eye moves closer to the central fixation point. (Figures by Johannes M. Zanker.)

GRADIENT PATTERNS AND BACKGROUND BRIGHTNESS CHANGES

A different group of static motion illusions are experienced in patterns that contain geometric arrangements of elements that feature luminance or color gradients, most famously illustrated by Akiyoshi Kitaoka in his rotating snakes illusion (Conway, Kitaoka, Yazdanbakhsh, Pack, & Livingstone, 2005; Kitaoka & Ashida, 2003), a set of concentric rings made from yellow-black-blue strips, which start moving vividly when the observer moves the eyes across the image but stand still when the observer steadily fixates a reference point. A substantially simplified, and as a result less powerful, version of an isolated rotating snake is illustrated in Figure IV.85-3a. Although the concentric arrangement of disks that are filled with a grey-level gradient only retains a semblance of Kitaoka's original design, at sufficient magnification and high contrast the circles of discs appear to spin around the center of the rings and most notably do so as well as the original design when the pattern is briefly flashed or switched on, or when the observer blinks. For comparison, an optimized version of the Pinna-Brelstaff illusion (Pinna & Brelstaff, 2000; see also chapter 74 in this volume), constructed from concentric sets of oriented Gabor patches (Gurnsey, Sally, Potechin, & Mancini, 2002), is shown in Figure IV.85-3b: when moving the eyes closer toward the fixation spot in the center, or further away, the observer perceives a very strong and puzzling sensation of rotating patches, despite the fact that the position of each patch within the circle remains constant. It is clear that these motion illusions based on systematically tilted elements of luminance or color contours rely on image displacements on a much larger scale than the micro-saccades that trigger motion illusions in phase-modulated gratings and also lead to coherent motion patterns on a larger scale. It is also worth noting that smooth gradients are not strictly necessary, as demonstrated in a variant of Figure IV.85-3b that is further reduced to sets of small, tilted-line segments (Gori & Hamburger, 2006) and the discontinuous patches in the rotating snakes. In the following, we show that the same computational modeling approach that was used in the previous section to analyze the outputs of motion detector networks for stimuli generating motion illusions can deliver satisfactory explanations as to why such illusions are induced by a variety of oriented gradient patterns. The focus here is on a version shown in Figure IV.85-3a, which shares a number of properties with the rotating snake, such as the coherent rotation perceived for large image shifts, but also can be extended into a stimulus that no longer requires any displacements. In such configurations, this stimulus is particularly instructive in terms of computational mechanisms that can be applied in similar ways, however, to other versions of this group of illusions.

The observation that sharp, temporal transients of the stimulus shown in Figure IV.85-3a, in the absence of any image displacement, elicit rather vivid motion illusions was the starting point to another stimulus configuration that generates a lasting, coherent, and strong motion sensation, which was dubbed the "spinning disk illusion" (Williams et al., 2005): when the brightness of the homogenous background of the disks is decreased or increased, the entire group of discs appears to rotate clockwise or counterclockwise, respectively, and does so even if the observer steadily fixates a reference point. It should be noted that this stimulus configuration, in terms of its fundamental properties, resembles a display used to demonstrate illusory motion reversals some 30 years ago (Anstis & Rogers, 1975; Gregory & Heard, 1983), as well as an inversion of foreground and background patterns discussed in the context of contrast asynchronies (Shapiro, Charles, & Shear-Heyman, 2005; Shapiro, Knight, & Lu, 2011). Although in this case the stimulus is not strictly static, because the background luminance is changing in time, it does not contain any motion energy because for each location luminance is either modulated in space but constant in time (pure spatial gradient), or it is modulated in time but constant in space (pure temporal changes). Most crucially, at any given location in space and time there is no joint modulation in *space and time* that is the defining element of *motion energy*. It is interesting to note, however, that both aspects are represented in the boundary regions between disks and background, albeit without overlap at any particular location.

So how does the computational model of cortical motion detector networks that is designed to extract spatiotemporal correlation (introduced in the previous section as 2DMD model) respond to such stimuli that contain no motion energy? A minimal version of this illusion is the spinning disk illusion (see Fig. IV.85-3a), shown in more detail in the top of Figure IV.85-4a: two disks filled with a grey-level gradient sloping in opposite directions are displayed in front of a background luminance that is increasing (left side) or decreasing (right side). The corresponding 2DMD output maps presented at the bottom of this figure show rings of strong motion responses, which outline the (static) disk in the motion signal maps in the boundary regions between the disks and their background. Most important, the disks boundaries are marked with green and red colors for opposite disk gradients, which correspond to rightward and leftward motion, respectively, and the colors are inverted for opposite directions of brightness change—in perfect correspondence to the perceived illusion. Note that the model response is limited to the boundary regions whereas the disks appear to move as a whole, suggesting a filling-in mechanism for these clearly outlined figures. This local motion response at the boundaries, in computational terms, is due to displacements of the contrast profiles at the edges of the disk resulting from the lateral inhibition that is needed to increase directional selectivity of the motion detectors by removing static signal components from their input (Zanker, 2007). These responses of the 2DMD model reflecting the spatial gradient-induced perception of illusory motion therefore can be regarded as a sufficient explanation of the illusion in terms of the cortical mechanisms that underlie the processing of visual motion information in the human brain. In this particular case, the model generates motion signals in the absence of motion energy in the physical stimulu as a byproduct of the neural filters used to construct the biologically plausible model.

Understanding the local processing mechanisms underlying the spinning discs illusion allows us to design a plethora of novel stimulus configurations that make patterns move in a paradoxical absence of displacements, just

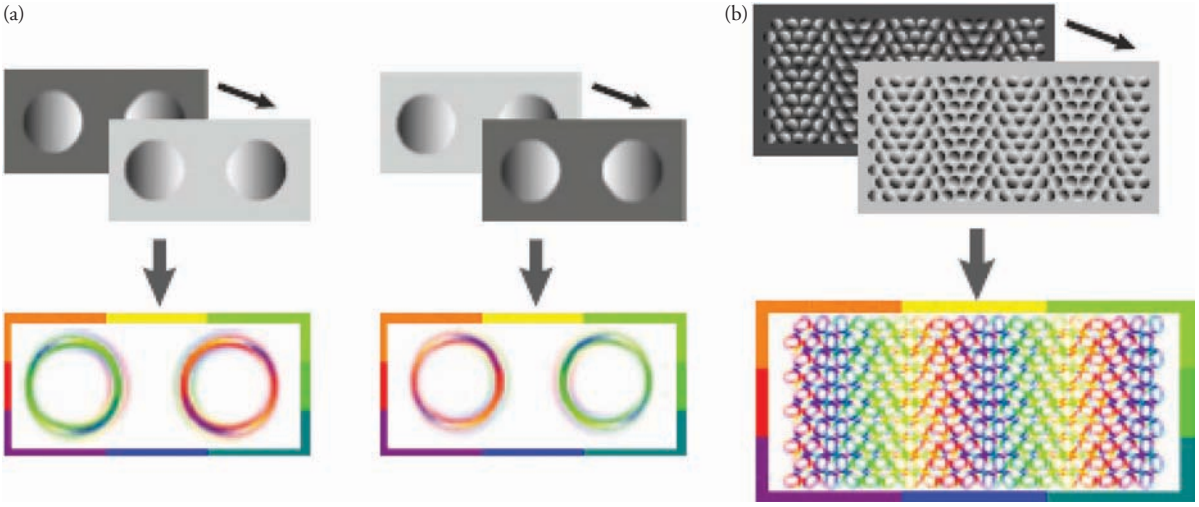


Figure IV.85-4. Modeling motion responses (from the 2DMD model) in the absence of motion energy for the spinning disk illusion: static disks filled with grey-level gradients are presented in front of a homogenous background with increasing or decreasing luminance. (a) Isolated disks with opposite horizontal gradients (top) generate motion responses (bottom panels, color code as illustrated in Fig. IV.85-2b) rightward (green) or leftward (red), depending on the direction of luminance change. (b) In an array of disks, a horizontal variation of gradient orientation (top) leads to a systematic variation of motion direction signaled by the 2DMD model (see color variation in bottom panel) and of perceived motion direction, which can be described as wave illusion. (Figures by Johannes M. Zanker.)

by manipulating the local orientation of gradient elements in a set of discs. This includes stimuli that are perceived as moving without any change of position, which therefore lend themselves to brain imaging studies that benefit from retinotopic stability (see Williams et al., 2005). The example illustrated in Figure IV.85-4b shows regular two-dimensional disk arrays in which the orientation of the grey- level gradient is constant for each set that shares the same horizontal position in the grid but varies according to a sine-wave function along the horizontal dimension of the grid elements—when this pattern is animated by oscillating the background luminance, observers report a strong and compelling sensation of a wave moving across the stimulus display, which is nicely reflected by the motion signal maps generated by the 2DMD model shown at the bottom of Figure IV.85-4b. More generally, flow-fields, as seen by an observer translating through, or rotating in, a natural environment can be constructed with sets of randomly distributing disks with radial or tangential orientation of the grey-level gradient and by modulating the background luminance. Some work in progress (Zanker, 2009) compares the optic flow measures from 2DMD output distributions for such flow-fields with the performance of human observers in detecting and discriminating such patterns. These experiments can demonstrate that the computational model captures the essence of the illusion and further specify the crucial stimulus parameters, and therefore allow us to assess the utility of such illusory stimuli as quantitative tools to study motion processing.

CONCLUSIONS

The purpose this chapter was to illustrate, however briefly, that behind the term "static motion illusions" we can discover a variety of unexpected percepts that can be traced back to rather different sources, such as frequent involuntary but small eye movements that usually go unnoticed or inherent properties of neural networks that can have surprising consequences. Understanding such phenomena requires the careful definition of stimulus conditions in terms of space and time and consideration of the visuo-motor interaction and the resulting space-time characteristics of the input to cortical processing networks, accompanied by quantitative modeling of the neural networks that generate (in this case, illusory) perception. In such an attempt to comprehend a plethora of puzzling phenomena, and to bring some sense of order into the jolly cottage industry of static and dynamic illusions, we experience an exciting partnership between disciplines: artists compose, scientists dissect, and the audience is surprised.

REFERENCES

Anstis, S. M., & Rogers, B. J. (1975). Illusory reversal of visual depth and movement during changes of contrast. *Vision Research*, *15*, 957–961.

Campbell, F. W., & Robson, J. G. (1958). Moving visual images produced by regular stationary patterns. *Nature*, *181*, 362.

Conway, B. R., Kitaoka, A., Yazdanbakhsh, A., Pack, C. C., & Livingstone, M. S. (2005). Neural basis for a powerful static motion illusion. *The Journal of Neuroscience*, *25*, 5651–5656.

Gori, S., & Hamburger, K. (2006). A new motion illusion: The rotating-tilted-lines illusion. *Perception*, *35*, 853–857.

Gregory, R. L. (1993). A comment: MacKay rays shimmer due to accommodation changes. *Proceedings of the Royal Society London, B*, *253*, 123.

Gregory, R. L., & Heard, P. (1983). Visual dissociations of movement, position, and stereo depth: Some phenomenal phenomena. *Quarterly Journal of Experimental Psychology*, *35A*, 217–237.

Gurnsey, R., Sally, S. L., Potechin, C., & Mancini, S. (2002). Optimising the Pinna-Brelstaff illusion. *Perception*, *31*, 1275–1280.

Kitaoka, A. (2002). *Trick eyes*. Tokyo: Kanzen.

Kitaoka, A., & Ashida, H. (2003). Phenomenal characteristics of the peripheral drift illusion. *Vision*, *15*, 261–262.

Leviant, I. (1996). Does "brain-power" make Enigma spin? *Proceedings of the Royal Society London, B*, *263*, 997–1001.

MacKay, D. M. (1957). Moving visual images produced by regular stationary patterns. *Nature*, *180*, 849–850.

Martinez-Conde, S., & Macknik, S. L., (2007, August). Windows on the mind. *Scientific American*, 56–63.

Pinna, B., & Brelstaff, G. J. (2000). A new visual illusion of relative motion. *Vision Research*, *40*, 2091–2096.

Shapiro, A. G., Charles, J. P., & Shear-Heyman, M. (2005). Visual illusions based on single-field contrast asynchronies. *Journal of Vision*, *5*, 764–782.

Shapiro, A. G., Knight, E. J., & Lu, Z.-L. (2011). A first- and second-order motion energy analysis of peripheral motion illusions leads to further evidence of feature blur in peripheral vision. *PLoS ONE*, *6*, e18719.

Troncoso, X. G., Macknik, S. L., Otero-Millan, J., & Martinez-Conde, S. (2008). Microsaccades drive illusory motion in the Enigma illusion. *Proceedings of the National Academy of Sciences USA*, *105*, 16033–16038.

Wade, N. J. (1982). *The art and science of visual illusions*. London: Routledge & Kegan Paul.

Williams, A. L., Zanker, J. M., & Ashida, H. (2005). Cortical activity during illusory motion sensations: The spinning disks illusion. *Perception*, *34*, 125.

Zanker, J. M. (2004). Looking at op art from a computational viewpoint, *Spatial Vision*, *17*, 75–94.

Zanker, J. M. (2007). A simple motion-detector model explains the spinning-disks illusion. *Perception*, *36*, 64.

Zanker, J. M. (2009). Optic flow perceived from illusory disk motion. *Perception*, *38*, 10.

Zanker, J. M., Doyle, M., & Walker, R. (2003). Gaze stability of observers watching op art pictures. *Perception*, *32*, 1037–1049.

Zanker, J. M., Hermens, F., & Walker, R. (2010). Quantifying and modeling the strength of motion illusions perceived in static patterns. *Journal of Vision*, *10*, 1–14.

Zanker, J. M., & Walker, R. (2004). A new look at op art: Towards a simple explanation of illusory motion, *Naturwissenschaften*, *91*, 149–156.

Zeki, S. (1994). The cortical Enigma: A reply to Professor Gregory. *Proceedings of the Royal Society London, B*, *257*, 243–245.

Zeki, S., Watson, J. D. G., & Frackowiak, R. S. J. (1993). Going beyond the information given: The relation of illusory visual motion to brain activity. *Proceedings of the Royal Society London, B*, *252*, 215–222.

PART V
Faces

Chapter 86
The Venus Effect

Marco Bertamini and Richard Latto

DEFINITION

The Venus effect arises in situations where an observer sees a scene in which another person (called a subject in this chapter), is present together with a mirror and the observer can see both the subject and the subject's image reflected in the mirror. When describing this scene, there is a strong tendency for the observer to say that the subject is looking at his or her own image in the mirror. Therefore the subject in the scene is judged to be able to see a reflection of his or her own face, even when this is not possible given the layout of the scene, as shown in Figure V.86-1 (Bertamini, Latto, & Spooner, 2003).

The definition in Bertamini et al. (2003) referred to how pictures are perceived. It specifically described the situation in which a subject and a mirror are not placed along the observer's line of sight (e.g., a Venus admiring herself in a small mirror). The authors suggested that observers are influenced by the fact that the reflection of the subject in the mirror is visible to them. This is a type of egocentric bias, in the sense that observers assume that everyone—including the subject—has access to the same information from the mirror as they do. Figure V.86-1 shows a typical layout with a subject and a small mirror. Note that the face is visible to the observer only if the observer moves behind the subject with respect to the mirror. This is because in order to share a similar view of what is reflected in a mirror it is necessary to share a similar line of sight. Conversely, consider what happens when a subject, a mirror, and an observer form a triangle, as in Figure V.86-1. It is possible to orient the mirror so that the observer sees a reflection of the subject's face in it. This creates a basic symmetry because the subject will now see a reflection of the observer's face in the mirror. Therefore, based on experience, observers should feel "observed" by the subject when the subject's face is visible to them in a small mirror. Surprisingly, this does not seem to happen. Very few people, for example, would say that Veronese's Venus in Figure V.86-2 is using the mirror to look out of the picture, or toward the observer.

TOILET OF VENUS

The name of the Venus effect derives from the numerous paintings in which a Venus is shown at her toilette. The subject is present in Greek and Roman art but is probably more famous in the paintings of the Italian Renaissance. An early and famous example is by Titian (1488–1576), now at the Hermitage (Poglayen-Neuwall, 1934). The theme is that of beauty, and Venus is often shown together with jewelry, flowers, and a small mirror. Because both Venus and the mirror are important items in the composition, they are usually placed so that they are easily visible, roughly at the same distance from the viewer and side by side. Moreover, the face of Venus is almost always reflected

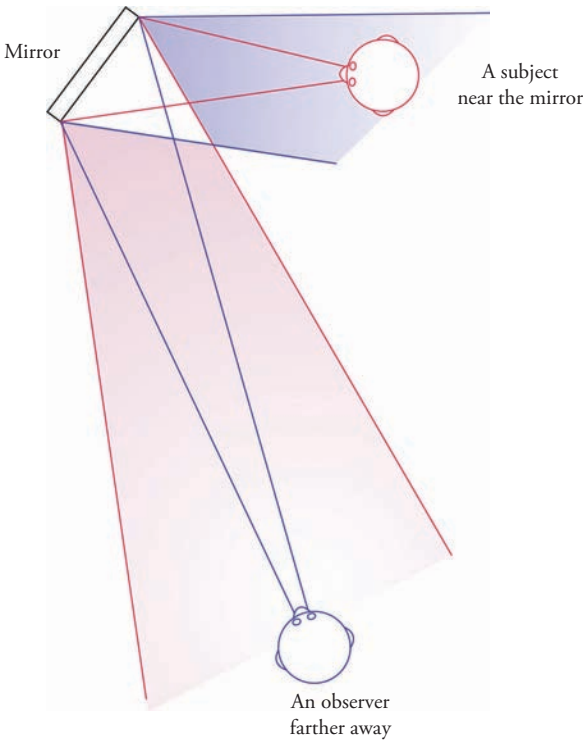


Figure V.86-1. This diagram shows a layout in which an observer (in blue) or a camera can see the face of a subject (in red) in the small mirror while standing in a different position. The blue shaded area is what is visible as reflected in the mirror for the observer, and the red shaded area is what is visible as reflected in the mirror for the subject. Note that the subject can see the observer's face but is not able to see his or her own face. (Image created by MB.)

in the mirror. In an analysis (Bertamini et al., 2003) of the paintings reproduced in the catalogue for Jonathan Miller's exhibition "On Reflection" at the National Gallery in London (Miller, 1998) we found 14 in which a person was visible at the same time as his or her own reflection in a small mirror. These were images likely to generate the Venus effect. The type of reflection, however, varied considerably, from a full face (Hans von Aachen, *Couple with Mirror*, c. 1596, Vienna, Kunsthistorisches Museum) to a profile facing in the same direction as Venus (Simon Vouet, *Allegory of Prudence*, c. 1645, Montpelier, Musée Fabre) to an opposite profile (Vasari, *Toilet of Venus*, 1558, Stuttgart, Staatgalerie). Despite all these variations, most viewers read the scene in a similar way, generally as a Venus intent on admiring herself in the mirror. The Venus effect thus allows artists great flexibility in how they represent a reflection without destroying the illusion of self-regard.

It is remarkable how much has been written about these famous works of art without any discussion of this basic issue of perception. Leaving aside the intentions of the artists, observers should be familiar enough with mirrors to understand that Venus is not looking at herself, given that the observer can see her face from a very different viewpoint. The phenomenon is not confined to the special case of Venus. One of the earliest examples of this kind of picture is *La dame à la licorne* (end of 15th century, Musée de Cluny, Musée du Moyen-Âge, Paris, France). The tapestry is complete with its six pieces. Five of them illustrate each of the five senses, and in the piece about "Sight," there is a unicorn with a woman holding a small mirror (Fig. V.86-2). To know whether this image generates a Venus effect, we must know how people interpret this scene. Fortunately, all we need to do is read the description of the Musée de Cluny (*The Lady and the Unicorn*, "Sight," n.d.): "in a familiar manner, the unicorn has placed its front paws on the Lady's knees and looks at its reflection in a mirror she holds." The Wikipedia entry on the tapestries also reports that "the unicorn … gazes at his reflection in the mirror" (*The Lady and the Unicorn*, 2011).

In Figure V.86-2 we provide an example from a Venus by Veronese. The pragmatic use of the mirror permitted to the artist by the Venus effect is highlighted by the description provided by the museum: "the inclusion of the mirror presents the opportunity to portray the goddess both full-face and in profile" (Joslyn Art Museum, n.d.). This shows how freely mirrors can be used without any clear reference to what Venus can and cannot see.

PICTURES, FILMS, AND REAL LIFE

The definition given at the beginning of this chapter, and also in Bertamini et al. (2003), describes a phenomenon that occurs in the perception of pictures. This is not, however, an effect confined to visual art. Indeed, a few observations from our own daily life illustrate that something similar also happens in situations where the person and the mirror are physically present. To test this, Bertamini, Lawson, Jones, and Winters (2010) placed a mannequin in a room and allowed people to look into the room from a side door. (A photograph of the observer's view is shown in Figure V.86-2.) The effect was just as strong under these conditions, with 73% of observers reporting that the mannequin was looking at herself in the mirror, a similar proportion to that found when viewing a painting of Venus at her toilette (Bertamini et al., 2003).

The situation where the Venus effect is most common is in film and television programs. As in the paintings of the toilet of Venus, it is often the case that an actor is looking at a mirror. Even when the viewer (i.e., the camera) is clearly not behind the actor and the viewer sees the actor's face in the mirror, the impression is not that of the actor looking toward the viewer or toward the camera. As with paintings, this phenomenon is convenient because a director can arrange actors and mirrors in many ways, and frame the face of the actor for the camera, without worrying about viewers being puzzled by an actor unable to see his or her own reflection. It should be noted that as the mirror gets larger, it becomes possible for both the actor and the camera to see the actor's face (see again Fig. V.86-1 and imagine a larger mirror surface, thus extending the space visible in the reflection). Nevertheless, the layout is often chosen to frame the image of the actor to appear in the center of the mirror for the camera, rather than for the actor (see the framing of the image in *Snow White*, Figure V.86-2 or the final scene of *Raging Bull* in which Robert De Niro is talking to himself in front of a large mirror).

Another important question is whether the presence of the visible reflection of the subject's face is necessary for the Venus effect. That is, do viewers say that the subject sees what they see? This would be an example of egocentrism.



Figure V.86-2. Tapestry: *La dame à la licorne* (end of 15th century). Detail. Musée de Cluny, Paris. Painting: Veronese (1528–1588), *Venus at Her Toilette*, c. 1582. Joslyn Museum, Omaha. Film: *Snow White and the Three Clowns* (1961), Twentieth Century Fox. Real life: A mannequin in a study (Bertamini et al., 2010). (Images used under academic fair use, and photograph taken by MB.)

However, evidence from Bertamini et al. (2010) suggests that the answer is no. As long as a subject is seen near a mirror and is looking toward it, observers tend to claim that the subject can see him or herself in the mirror even though the reflection from the mirror does not show any face.

MIRRORS AS FAMILIAR OBJECTS

Many simple things in life are appreciated only when they are no longer around. It is difficult, though, to imagine a world without mirrors. Their existence is a significant part of our life, and it is almost impossible in an industrialized society to spend a single day without seeing some reflection in a mirror and see our own image reflected.

A New York artist, Anita Sto (2012), has recently performed an experiment in which she avoided seeing herself in a mirror or even in a photograph for an entire year (November 11, 2010 to November 11, 2011). In the process, she discovered how difficult this task is and how much of a difference mirrors make to her life. We make this point about the ubiquity of mirrors because from such vast experience people might be expected to have acquired accurate implicit and explicit knowledge about how mirrors work: for instance the fact that we can see our own reflection only when we are in front of a mirror and not when we are standing to its left or right. Yet several studies have shown that this is precisely the kind of knowledge most people do not have. For example, when asked where they need to move to in order to be able to see themselves, many people expect to be able to see themselves without being in front of the mirror (Bertamini, 2010; Bertamini et al., 2003). We call this a "from-the-side" error. Experience does not eliminate the problem, and in a way experience may make people overreliant on the general idea that mirrors show what is within a cone in front of them, as shown by the fact that the from-the-side error does not exist in primary schoolchildren (Bertamini & Wynne, 2009).

ARTISTS' EXPLOITATION OF THE VENUS EFFECT

The Venus effect differs from many of the other illusions described in this book. Most of these persist even when their nature is explained to the observer. On the other hand, once the Venus effect is brought to people's attention, it becomes relatively easy to spot examples where the subject is unable to see him or herself in the mirror. People can learn to reason about the effect. However, this most likely involves an overt analysis of the image, rather in the way representational artists learn techniques to enable them to work out how to project a three-dimensional scene onto a flat surface in order to paint it. Despite this training, the artist continues to *perceive* the world as we do, with distant objects scaled up to reflect their actual size, partially occluded objects seen as complete, and so forth. In the same way, our perceptual systems never fully come to grips with how mirrors work, however much we know about the optics of reflection. This is why artists of all kinds can play with the reproduction of mirrors and their reflections, sometimes using impossible reflections, without the image immediately looking wrong. Even in extreme examples, like Magritte's painting *La Reproduction interdite* (Figure V.86-3), it takes us a moment to identify, if indeed we do at all, the several things wrong with the image.

The presence of mistakes in the reproduction of mirror reflections is interesting, but it can give the wrong impression in relation to the Venus effect. The Venus effect does not imply anything wrong with the picture and, as discussed, takes place even with photographs, film, and real people in front of real mirrors. The Venus effect is a misunderstanding of what another person is able to see in a mirror. In this sense it can exist independently of what the painter, director, or photographer intended to represent.

KNOWLEDGE ABOUT VIEWPOINT

The experimental study of the understanding of perspective in children originated with Piaget. In his "three mountains" task, a child was asked what a doll could see of a landscape. The task requires an appreciation that what is visible to different people varies with their viewpoint (Piaget & Inhelder, 1967). Until the age of four, children struggle to distinguish between their own view and that of the doll (egocentrism). Later research has shown that children's perspective-taking ability is affected by a variety of situational variables and that by the age of five most children will understand that an object appears different to observers who view it from different sides (Flavell, Flavell, Green, & Wilcox, 1981). On the other hand, some egocentric biases do also exist in adults (Epley, Morewedge, & Keysar, 2004). The Venus effect is an extreme example of a problem with reasoning about perspective. Our cognitive processes do not seem to include the facility to incorporate fully the way mirrors work, and so, like Piaget's infants, we fall back on our intuitions. These tell us that mirrors reveal a large spatial area, which is a copy of what is located in front of the mirror itself, within a cone that extends both left and right (Bertamini et al., 2010; Bertamini, 2014).

Relying on our memory of objects and their layout, we cannot easily answer questions about viewpoint (what is visible from where) because the viewpoint is nothing more than an accidental aspect of the visual experience. To understand what we mean by "accidental," consider for example that the wheel of a bicycle is an ellipse from almost all viewpoints. For the visual system, that is a distraction (an accident) of no consequence, and the wheel is always perceived and remembered as circular. These processes are known as constancy mechanisms (size constancy, shape constancy, and so on).

THE PROJECTION ON A MIRROR

We have described how people find it difficult to judge what is made visible by means of a mirror to themselves or other people. This is a problem with viewpoint-dependent information. But a mirror is also an object on the wall. Could people learn what is "painted" on that object? In other words, could people treat mirrors as pictures and, in doing so, learn about what is shown in the picture? The answer is people find it extremely difficult to look at a mirror as if

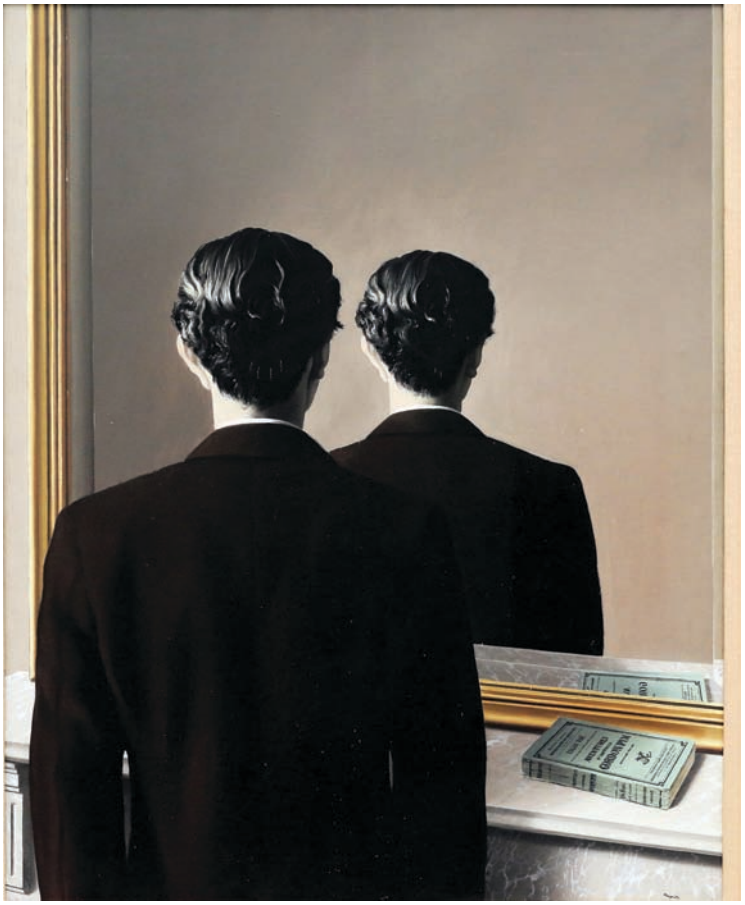


Figure V.86-3. Magritte (1937), *La reproduction interdite*. Museum Boijmans Van Beuningen, Rotterdam, The Netherlands. As well as the anomalous reflection of the subject's back, the reflection should be rotated slightly, not seen square on. In this and many other artistic examples (e.g., Manet's *Le Bar aux Folies-Bergère*), the freedom with which the artist has created the reflection is not obvious until pointed out. (Image used under academic fair use.)

it were a picture. The best example is when observers are asked about the size of the image projected on a mirror.

Gombrich (1960) suggested a demonstration that everyone can do in their own homes. Look at a fogged-up mirror in your bathroom and clear just enough space to see your own face. You will be surprised by how small this area is. It is in fact half of the height of your face (a quarter of the area). When told this, most people find it hard to accept that the image of their own face on the glass surface is only half the height of their face. In addition, they find it even harder to accept that this is true whatever their distance from the mirror (Bertamini & Parks, 2005). Thus the problem of trying to judge size of projections persists even when the image is right in front of our eyes (Bertamini Lawson, R., & Liu, 2008; Lawson, Bertamini, & Liu, 2007). Readers are urged to test this at home using a felt-tip pen and closing one eye (the presence of two eyes is a minor complication), rather than taking our word for it.

NAIVE PHYSICS

Systematic differences between people's intuitive understanding of the physical world and what they should have learned from experience have been well documented. This area of research is known as "naive physics" (Proffitt, 1999). For example, people have intuitions about how objects move and fall, and they expect an object moving along a curved path to continue along such a path when released (McCloskey, 1983). The Venus effect can be seen as another example of an incorrect intuition about how the world works. The mistakes people make have multiple sources. In part they may originate from problems with reasoning, but they may also arise as a consequence of characteristics of perception. As mentioned earlier, the visual system has evolved to extract information about solid objects and discard accidental information such as changes in viewpoint. This factor helps to explain why it is hard to judge the properties of the projected image on a mirror. It is also worth remembering that although mirrors are very familiar to us today, they were not present in the environment in which we evolved.

REFERENCES

Bertamini, M. (2010). Mirrors. *The Psychologist*, *23*, 112–114.

Bertamini, M. (2014). Understanding what is visible in a mirror or through a window before and after updating the position of an object. *Frontiers in Human Neuroscience*. 10.3389/fnhum.2014.00476

Bertamini, M., Latto, R., & Spooner, A. (2003). The Venus effect: People's understanding of mirror reflections in paintings. *Perception*, *32*, 593–599.

Bertamini, M., Lawson, R., Jones, L., & Winters, M. (2010). The Venus effect in real life and in photographs. *Attention, Perception, & Psychophysics*, *72*, 1948–1964.

Bertamini, M., Lawson, R., & Liu, D. (2008). Understanding 2D projections on mirrors and on windows. *Spatial Vision*, *21*, 273–289.

Bertamini, M., & Parks, T. E. (2005). On what people know about images on mirrors. *Cognition*, *98*, 85–104.

Bertamini, M., Spooner, A., & Hecht, H. (2003). Naive optics: Predicting and perceiving reflections in mirrors. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(5), 982–1002.

Bertamini, M., & Wynne, L. A. (2009). The tendency to overestimate what is visible in a planar mirror amongst adults and children. *European Journal of Cognitive Psychology*, *22*, 516–528.

Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, *40*(6), 760–768.

Flavell, J. H., Flavell, E. R., Green, F. L., & Wilcox, S. A. (1981). The development of three spatial perspective-taking rules. *Child Development*, *52*, 356–358.

Gombrich, E. H. (1960). *Art and illusion.* Oxford: Phaidon Press.

Joslyn Art Museum. (n.d.). *Venus at her toilette*. http://www.joslyn.org/collections-and-exhibitions/permanent-collections/european/veronese-venus-at-her-toilette/

Keil, F. C. (2003). Folkscience: Coarse interpretations of a complex reality. *Trends in Cognitive Science*, *7*, 368–373.

Lawson, R., Bertamini, M., & Liu, D. (2007). Overestimation of the projected size of objects on the surface of mirrors and windows. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 1027–1044.

McCloskey, M. (1983). Intuitive physics. *Scientific American*, *248*(4), 114–122.

Miller, J. (1998). *On reflection*. London: National Gallery Publications.

Piaget, J., & Inhelder, B. (1967). *The child's conception of space*. New York: Norton.

Poglayen-Neuwall, S. (1934). Titian's pictures of the toilet of Venus and their copies. *Art Bulletin*, *16*(4), 358–384.

Proffitt, D. R. (1999). Naive physics. In R. Wilson & F. Keil (Eds.), *The MIT encyclopedia of the cognitive sciences*. Cambridge, MA: MIT Press.

Sto, A. (2012). *Everything begins with an object in the real world: One year performance piece.* New York: Hart Repro.

The lady and the unicorn. (2011). Wikipedia. http://en.wikipedia.org/wiki/The_Lady_and_the_Unicorn

The lady and the unicorn. "Sight." (n.d.) http://www.musee-moyenage.fr/collection/oeuvre/la-dame-a-la-licorne.html

Chapter 87
The Hollow-Mask Illusion and Variations

Thomas V. Papathomas

INTRODUCTION—HISTORY

The hollow-mask, or hollow-face, illusion is perhaps the best-known *depth inversion illusion*: a three-dimensional (3D) object is perceived in reverse depth. It is a prime example of a class of objects that exhibit depth inversion. In the strict version of *global* depth inversion, for any pair of points A and B on the surface of the object, if A is physically further/closer than B, A is perceived to be closer/further than B; in this case, convex parts of the object appear as concave and vice versa. There are cases in which this relationship holds only locally; that is, depth is inverted only for certain parts of the object, whereas the rest of the parts appear in veridical depth. The chapters on reverse-perspective illusions and on the Ames window in this volume present other examples of depth inversion illusions.

The hollow-mask illusion is relatively well known (Gregory, 1970, 1997): a hollow mask appears as a normal convex face as long as it is seen from a sufficiently large distance. Brewster (1826/1983) devised one of the first descriptions of the hollow-mask illusion (see also Wade, Campbell, Ross, & Lingelbach, 2010). According to Wheatstone (1838), the first to report the "apparent conversion of a cameo [design raised above the background of a gemstone] into an intaglio [design carved into a gemstone], and of an intaglio into a cameo" was "Professor Gmelin, of Wurtemberg, [who] published a paper on the same subject in the Philosophical Transactions for 1745" (p. 383). In ancient civilizations, sculptors who used molds to design bronze statues must have noticed some form of this illusion on the discarded molds, since the hollow-face illusion generalizes to other parts of the body, such as the torso (Papathomas, de Heer, Zhuang, Grace, & Bunkin, 2012). Nevertheless, I know of no references to the hollow-mask illusion that predate Gmelin's 1745 publication. Gregory (1970, page 128) speculates that "The effect may have been the basis of ancient temple miracles: statues moving as though alive, following each devotee with their hollow heads as he moves in the gloom of the sacred place."

Depth inversion is accompanied by a stunning percept of illusory motion: namely, as a viewer moves in front of a hollow mask, the face—which is perceived as convex—appears to follow the viewer (Gregory, 1970; Papathomas, 2007; Rogers & Gyani, 2010). The illusion is generally attributed to the top-down influences of familiarity and expectation (Gregory, 1970, 1997; Yellott, 1981). After all, we have a lifelong experience with convex faces. The fact that the illusion is significantly weaker when the hollow mask is turned upside-down (Hill & Bruce, 1993; Hill & Johnston, 2007; Papathomas & Bono, 2004) offers strong evidence for such high-level, schema-driven influences (i.e., we are more familiar with upright than inverted faces).

EXPLANATION

Why does the mask "move" as we move in front of it? One possible explanation for the apparent illusory motion is provided by Figure V.87-1 (Papathomas, 2007); it provides a top view of a hollow mask (solid black lines), as seen by a viewer who starts at position 1 to the left of the mask and moves to position 2 while viewing the mask. The percept of a (illusory) convex face at position 1 is indicated by dashed lines and blue fill. The dashed blue lines of sight

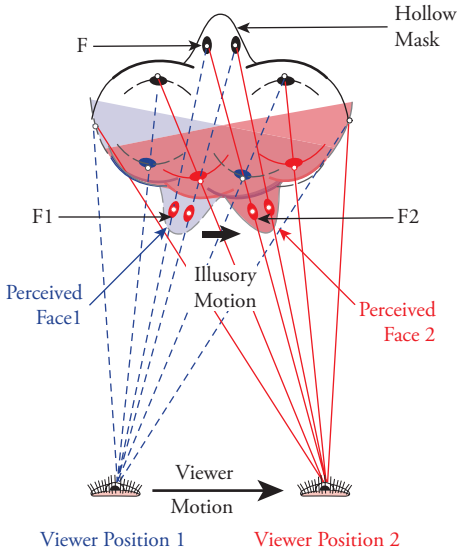


Figure V.87-1. A possible explanation for the hollow-mask illusion. (Modified from Figure 3 in Papathomas, 2007, © Brill Publishers.) (Figure taken from Papathomas (2007) with permission.)

for *physical* key features, such as head edges, eyeballs, and nostrils, constrain the positions of the corresponding *perceived* features. Any feature, such as the left nostril F, must lie along its own corresponding line of sight that is anchored at the physical feature; F is perceived as the right nostril located at F1. When the viewer moves to position 2, each line of sight (shown solid) is still anchored at a *physical* point, such as F, on the hollow mask. If the viewer maintains the illusion, the corresponding *perceived* feature must lie on the new line of sight and thus is perceived at position F2. This is why the feature F moves from F1 to F2 as the viewer moves from position 1 to 2. This happens with every feature on the mask. The position of the convex mask is now indicated by solid grey lines (for the outline) and red fill; the facial features are shown by red solid lines. This gives rise to the compelling illusory motion in which the mask appears to "follow" the viewer, as shown in Figure V.87-1.



Video V.87-1. Video by Thomas V. Papathomas.

PERCEIVING A HOLLOW MASK TO ROTATE IN THE OPPOSITE DIRECTION

An example of a hollow mask is shown in Figure V.87-2, which shows five frames of an animation sequence of a rotating mask. If the mask is viewed from above, it rotates in the counterclockwise (CCW) direction, as indicated by the white arrows. When frames 1 and 2 are viewed statically, it is extremely hard to perceive either as a hollow mask; instead, they appear as convex faces. Even the image of frame 3, in which the convex surfaces of the left temple and cheek start to occlude the corresponding concave surfaces, appears as a convex face, unless the viewer pays attention to the self-occlusion. The concavity of the mask is evident in frames 4 and 5, in which the self-occlusion is clearly visible. It is clear from the first three frames that if the viewer maintains the hollow-mask illusion, then the illusory convex face appears to rotate clockwise (CW; as seen from above), opposite to the real direction of rotation. An example is provided on Video VI.87-1 on the Companion Website.

FACTORS THAT STRENGTHEN OR WEAKEN THE ILLUSION

Generally, the bottom-up signals of binocular disparity and motion parallax provide cues to recover the veridical 3D depth relationships, and so they counter, or weaken, the illusion. The extraretinal bottom-up signals of vergence and accommodation may play a similar role, but their influence has not been established. Cues that strengthen the illusion include the top-down influence of familiarity and expectation. Finally, cues that can be manipulated to favor or oppose the illusion include shadows and painted features. Based on these factors, one can make the following predictions regarding what effects can be expected as a result of stimulus manipulations and viewing conditions.

1. Binocular viewing: This weakens the illusion because it adds the important signal of disparity (Kaur, Papathomas, & DeCarlo, 2000; Papathomas & Bono, 2004; Sherman, Papathomas, Jain, & Keane, 2011). This is why people are encouraged to view the hollow mask monocularly for a stronger illusion.
2. Moving/stationary viewer: When the viewer moves, he or she obtains motion parallax signals that are important for recovering the 3D structure of the stimulus. Thus the illusion is stronger for a stationary viewer (Papathomas et al., 2010; Sherman et al., 2011).
3. Viewing distance: Both binocular disparity and motion parallax signals become weaker at large distances. Thus the illusion gets stronger with increasing viewing distance (Gregory, 1970; Papathomas & Bono, 2004).
4. The inversion effect: As mentioned, since we are more familiar with upright than upside-down faces, the illusion is weaker for a hollow mask that is shown

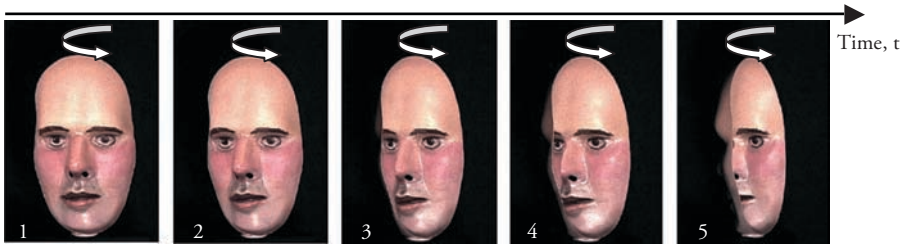


Figure V.87-2. A hollow mask rotating CCW, as seen from above (white arrows). Top-down processes (familiarity) cause it to be perceived as a convex face. This depth inversion is accompanied by illusory motion: the mask appears to rotate CW, as seen in frames 1 to 3. The illusion disappears when parts of the mask occlude other parts. The onset of self-occlusion is shown in frame 3; it is evident in frames 4 and 5. (Figure by Thomas V. Papathomas.)

upside-down (Hill & Bruce, 1993, 1994; Papathomas & Bono, 2004).

5. Shadows: When a hollow mask is lit from below, the resulting pattern of shadows resembles the pattern of shadows that are produced by a convex mask lit from above. Thus when we light the hollow mask from below or above, it results in a stronger or weaker illusion, respectively (Hill & Bruce, 1993).
6. Painted random features: When a random-dot texture is added to an unpainted hollow mask, resulting in a "dotted mask," the dots provide depth "anchors" that help recover the true 3D shape of the hollow mask; indeed, such a hollow mask elicits a weaker illusion as compared to an unpainted hollow mask (Papathomas et al., 2010).
7. Painted facial features: One would expect to obtain a stronger illusion for a hollow mask that is painted with realistic facial features (such as the one shown in Figure V.87-2) than for an unpainted hollow mask, based on familiarity, because the painted mask is closer to the faces that we encounter in everyday life. However, when this hypothesis is tested experimentally, there is a trend toward this direction but the difference is not significant (Kaur et al., 2000; Papathomas & DeCarlo, 2013). One possible reason for the lack of a significant difference can be that these painted features, although providing a more realistic stimulus for a stronger illusion, also provide depth "anchors" (see number 6 in this list) that contribute to a weaker illusion by providing depth signals that help recover the true 3D shape of the hollow mask.

DEVELOPMENTAL ISSUES

At what age do humans obtain the hollow-mask illusion? There are conflicting reports on this issue. Tsuruhara et al. (2011) examined whether infants perceive a rotating hollow mask displayed on a screen as convex using preferential looking and habituation paradigms and found no evidence that five- to seven-month-old infants obtain the illusion. On the other hand, Corrow, Granrud, Mathison, and Yonas (2011) used a preferential reaching paradigm with a physical mask and obtained evidence that six-month-old infants perceive the hollow-face illusion monocularly but not binocularly, in agreement with the expected trends. The different testing methods can explain the discrepancy in the findings.

SPECIAL POPULATIONS—EFFECTS OF DRUGS AND ALCOHOL

It has been well documented that, generally, people with schizophrenia are less susceptible to the 3D hollow-mask illusion than healthy controls of comparable age, gender, and educational level (Emrich, Leweke, & Schneider, 1997; Koethe et al., 2006; Schneider, Leweke, Sternemann, et al., 1996; Schneider et al., 2002; Keane et al., 2013). Dima et al. (2009) conducted a functional magnetic resonance imaging study and applied dynamic causal modeling to explain differences between patients and controls. They concluded that there was a strengthening of bottom-up processes and a weakening of top-down influences in patients as compared to controls, consistent with the hypothesis proposed at the beginning of this chapter.

It has been observed that administering cannabis to viewers weakens the illusion. The substance reaches its maximum effect within 3 hours after administration and maintains a residual influence on the strength of the illusion even days later (Emrich et al., 1997; Emrich et al., 1991). Regular cannabis users obtain a weaker illusion (Semple, Ramsden, & McIntosh, 2003). With respect to alcohol, the illusion is significantly weaker for highly intoxicated patients and patients with alcohol withdrawal symptoms (Schneider et al., 1998). As Brian Keane observed upon reviewing this chapter, interestingly, all three of these (schizophrenia, alcohol, and cannabis) are associated with visual hallucinations (including extreme alcohol withdrawal). Finally, sleep deprivation results in weakening of the hollow-mask illusion (Sternemann et al., 1997). This evidence indicates that pathological and drug-induced conditions elicit cognitive impairments that inhibit proper top-down influences—such as the generation of perceptual hypotheses—on processing of visual input.

VARIANTS OF THE ILLUSION

The hollow-mask illusion is the best known among the depth-inversion illusions. It is partly due to high-level influences such as familiarity with convex faces, as well as to the tendency of humans to perceive objects as convex (Langer & Bülthoff, 2001; Sherman et al., 2011); after all, most objects we deal with in daily life are convex (rocks, animals, body parts, fruits, tools, etc.). Thus one expects that this type of depth inversion must also be obtained with concave objects that are less familiar. This is indeed the case, as illustrated by Hill and Bruce (1994) with a "hollow potato." In fact, the illusion is strongest for an upright hollow human mask, and it is equally weak for an upside-down hollow human mask and a hollow potato, providing further evidence for high-level influences.

An interesting question is: If we add 3D objects to the hollow mask (such as a smoking pipe or a nose ring), what will be the percept when we rotate the mask? The result is a compelling illusion in its own right. The viewer perceives these objects rotating in the opposite direction to that of the mask. Examples of such illusions are shown in Figure V.87-3 (upright) and Figure V.87-4 (sidewise mask). Figure V.87-3 shows three frames of an animation sequence: an upright hollow mask with 3D spherical eyeballs stuck on the whites of the eyes and a nose ring passed through the nostrils.

The mask rotates CW as seen from above; however, if one views it from the front, it appears to turn CCW. Namely, as we go from frame 1 to 3, the face appears to turn to its left (solid arrows). However, the nose ring rotates in the opposite direction (dotted arrows). Less obvious is the rolling of the eyeballs; it is more visible in Figure V.87-4. What is remarkable is that, even though the entire object comprising the mask, nose ring, and eyeballs is rigid, it appears as nonrigid when it rotates, because some of its parts rotate in the opposite direction to that of the face.

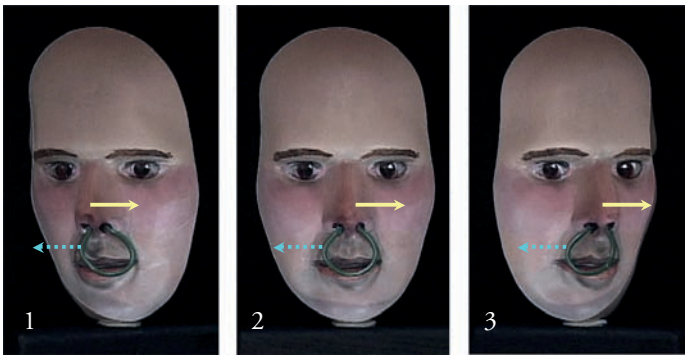


Figure V.87-3. A variation on the hollow-mask illusion. A nose ring and two spherical volumetric eyeballs have been added to the mask. Even though the entire mask/nose-ring/eyeballs object is rigid, it appears as nonrigid when it rotates. The nose ring and eyeballs appear to rotate in the opposite direction to that of the face. See http://www.youtube.com/watch?v=0QgoX78q-0Y. (Figure by Thomas V. Papathomas.)

Figure V.87-4 shows four frames of a sequence in which the same hollow mask, in a "reclining" posture, rotates CW (if seen from above; top thick arrow) but appears to turn CCW. Notice how the nose ring rotates in the opposite direction. In this figure, the rolling of the eyeballs (white arrows) is much clearer than in Fig. V.87-3.

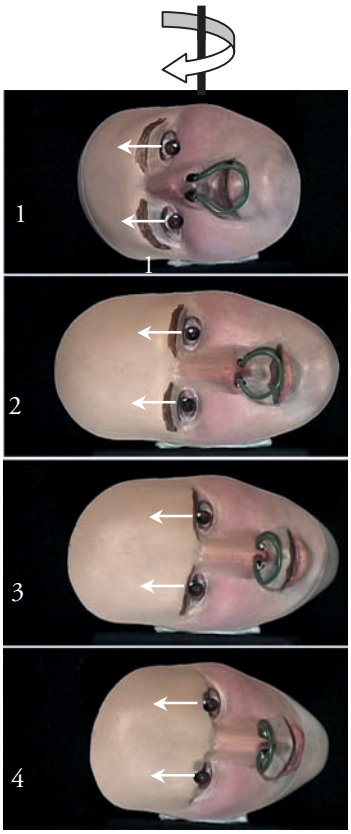


Figure V.87-4. Four frames in a sequence with a "reclining" hollow mask rotated CW as seen from above (top thick arrow). The mask appears to rotate CCW, but the nose ring and the eyeballs rotate CW, as indicated by the white arrows. Rolling of the eyeballs on the whites of the eyes is much more evident than in Figure V.87-3. (Figure by Thomas V. Papathomas.)

Why do we perceive a nonrigid object, even if the object is actually rigid? The reason is, if we add an external 3D object, such as the nose ring, it is highly unlikely that it will be assigned an illusory depth; hence it will not be subjected to the illusory motion that the mask is subjected to, and it will rotate in the veridical direction, opposite to that of the mask. This interesting motion illusion can be seen in Video V.87-2 on the Companion Website.

COMPUTER MODELING

DeCarlo and Metaxas (1996) have developed an algorithm that uses a two-dimensional image of a face to fit a 3D model for that face. Their scheme uses biometric parameters based on facial image features, such as the nose length and width, the interpupilary distance, the width of the eyes, the overall height and width of the face, and so on to estimate parameters that best fit a stored prototype. When their program is provided an animation sequence of a face, the algorithm tracks the face features and obtains a complete description of the motion trajectory. We provided an animated sequence of a painted upright concave rotating mask as input, and the algorithm estimated the mask's 3D shape and trajectory. The algorithm was susceptible to the hollow-mask illusion, just like humans; it perceived a convex face rotating in the opposite direction than that of the hollow mask (Kaur et al., 2000; Papathomas & DeCarlo, 2013). Of course DeCarlo and Metaxas's algorithm is endowed—just like humans are—with a top-down schema (prototype) of a convex 3D face. It is precisely the existence of this convex face schema that recovers the inverted-depth face structure that, in turn, forces the false interpretation of motion signals. This is not evidence that this is what happens with humans. Nevertheless, the susceptibility of the algorithm to the hollow-mask illusion provides some support for the explanation offered previously (Papathomas, 2007).

SIGNIFICANCE

The most important advantage of the hollow-mask illusion is that it can be used as a stimulus in experiments that enable us to infer what prior assumptions are used



Video V.87-2. Video by Thomas V. Papathomas.

by the visual system to construct a 3D representation of the world by pitting powerful data-driven signals against schema-driven influences. Like reverse-perspective stimuli (Chapter 30), it can also be used to study how we recover the depth relationships of objects in our environment as we move about. Because of differences in perception between patients and controls, it can also inform us about clinical disorders such as prosopagnosia, schizophrenia, or other problems involving psychosis.

Another important aspect of this illusion is the fact that it employs physical 3D stimuli. This aspect makes it much more relevant for studying visual mechanisms under natural conditions, as opposed to viewing planar (two-dimensional) stimuli. As Gregory (1994) put it,

> No perception in laboratory or practical class is complete without a hollow mask! But isn't it odd that James leaves out its apparent rotation following the observer wherever he moves at twice the speed he moves? This, indeed, shows the importance of having actual three-dimensional objects and not only pictures for appreciating perception. (p. 183)

The explanation proposed here assumes that the visual system uses a 3D representation of the environment, as well as vestibular and proprioceptive inputs, to predict the retinal flow under self-motion (Papathomas, 2007). The motion of external objects is determined by the degree to which the predicted retinal flow differs from the actual retinal flow. In addition to providing a rich area for studying perceptual questions, the hollow-mask illusion also serves as a vehicle for investigating issues in developmental and clinical psychology, as well as for building computer vision models with human relevance.

ACKNOWLEDGMENTS

Thanks to Tom Grace, Marcel de Heer, Brian Keane, Steve Silverstein, Robert Bunkin, Michael Cooper, Tim Maslyn, Akos Fehér, Anshul Jain, Rob Kohr, Alexander Papathomas, Hristiyan Kourtev, Lisa Bono, Alexandra Sherman, Anna Zalokostas, Vanja Vlajnic, James Hughes, Jordan Ash, Michael Marks, Yushi Wang, Deepthi Mikkilineni, Mark Theiler, Gyuri Schiff, Andrew Ng, Arielle Yeshua, and Joshua Dobias for their help in building the stimuli, designing/administering experiments with hollow masks, exploring variants of the hollow-mask illusion, and studying differences between schizophrenia patients and controls. Special thanks to Brian Keane for valuable suggestions on the chapter and Nicholas Wade for historical notes.

REFERENCES

Brewster, D. (1983). On the optical illusion of the conversion of cameos into intaglios, with an account of other analogous phenomena. In N. Wade (Ed.), *Brewster and Wheatstone on vision* (pp. 56–64). London: Academic Press. (Original work published 1826)

Corrow, S., Granrud, C. E., Mathison, J., & Yonas, A. (2011). Six-month-old infants perceive the hollow-face illusion. *Perception*, *40*, 1376–1383.

DeCarlo, D., & Metaxas, D. (1996). The integration of optical flow and deformable models with applications to human face shape and motion estimation. In *Proceedings 1996 IEEE Computer Society conference on computer vision and pattern recognition* (pp. 231–238). Los Alamitos, CA: IEEE Computer Society Press.

Dima, D., Roiser, J. P., Dietrich, D. E., Bonnemann, C., Lanfermann, H., Emrich, H. M., & Dillo, W. (2009). Understanding why patients with schizophrenia do not perceive the hollow-mask illusion using dynamic causal modeling. *NeuroImage*, *46*, 1180–1186.

Emrich, H. M., Leweke, F. M., & Schneider, U. (1997). Towards a cannabinoid hypothesis of schizophrenia: Cognitive impairments due to a dysregulation of the endogenous cannabinoid system. *Pharmacology, Biochemistry, and Behavior*, *56*, 803–807.

Emrich, H. M., Weber, M. M., Wendl, A., Zihl, J., Von Meyer, L., & Hanisch, W. (1991). Reduced binocular depth inversion as an indicator of cannabis-induced censorship impairment. *Pharmacology, Biochemistry, and Behavior*, *40*, 689–690.

Gregory, R. L. (1970). *The intelligent eye*. New York: McGraw-Hill.

Gregory, R. L. (1994). *Even odder perceptions*. London: Routledge.

Gregory, R. L. (1997). Knowledge in perception and illusion. *Philosophical Transactions of the Royal Society of London, B*, *352*, 1121–1128.

Hill, H., & Bruce, V. (1993). Independent effects of lighting, orientation, and stereopsis on the hollow-face illusion. *Perception*, *22*(8), 887–897.

Hill, H., & Bruce, V. (1994). A comparison between the hollow-face and "hollow-potato" illusions. *Perception, 23*, 1335–1337.

Hill, H., & Johnston, A. (2007). The hollow-face illusion: Object-specific knowledge, general assumptions or properties of the stimulus? *Perception, 36*, 199–223.

Kaur, M., Papathomas, T. V., & DeCarlo, D. (2000). Schema- and data-driven influences in the hollow-face illusion: Experiments and model. *Investigative Ophthalmology and Visual Science*, *41*(4), p. S224.

Keane B. P., Silverstein, S. M., Wang, Y., & Papathomas T. V., (2013). Reduced depth inversion illusions in schizophrenia are state-specific and occur for multiple object types and viewing conditions. *Journal of Abnormal Psychology, 122*(2), 506-512.

Koethe, D., Gerth, C. W., Neatby, M. A., Haensel, A., Thies, M., Schneider, U., . . . Leweke, F. M. (2006). Disturbances of visual information processing in early states of psychosis and experimental delta-9-tetrahydrocannabinol altered states of consciousness. *Schizophrenia Research, 88*, 142–150.

Langer, M. S., & Bülthoff, H. H. (2001). A prior for global convexity in local shape-from-shading. *Perception*, *30*, 403–410.

Papathomas, T. V. (2007). Art pieces that "move" in our minds—An explanation of illusory motion based on depth reversal. *Spatial Vision*, *21*, 79–95.

Papathomas, T. V., & Bono, L. (2004). Experiments with a hollow mask and a reverspective: Top-down influences in the inversion effect for 3-D stimuli. *Perception, 33*, 1129–1138.

Papathomas, T. V., & DeCarlo, D. (2013). Schema-driven influences in recovering 3-D shape from motion in human and computer vision. In S. Dickinson & Z. Pizlo (Eds.), *Shape perception in human and computer vision*. pp. 413–427, New York: Springer.

Papathomas, T. V., de Heer, M., Zhuang, X., Grace, T., & Bunkin, R., (2012). "Exorcist illusion": Twisting necks in the hollow-face and hollow-torso illusions. *i-Perception, 3*(10), 778–782.

Papathomas, T. V., Marks, M., Silverstein, S. M., Keane, B. P., Sinha, P., & Pavlovic, V. (2010). The role of pictorial cues in the hollow-face illusion: Binocular disparity versus motion parallax. *Perception*, *39*(ECVP Abstract Suppl.), 107.

Rogers, B. J., & Gyani, A. (2010). Binocular disparities, motion parallax, and geometric perspective in Patrick Hughes's "reverspectives": Theoretical analysis and empirical findings. *Perception, 39*, 330–348.

Schneider, U., Borsutzky. M., Seifert, J., Leweke, F. M., Huber, T. J., Rollnik, J. D., & Emrich, H. M. (2002). Reduced binocular depth inversion in schizophrenic patients. *Schizophrenia Research*, *53*, 101–108.

Schneider, U., Dietrich, D. E., Sternemann, U., Seeland, I., Gielsdorf, D., Huber, T. J., . . . Emrich, H. M. (1998). Reduced binocular depth inversion in patients with alcoholism. *Alcohol and Alcoholism*, *33*, 168–172.

Schneider, U., Leweke, F. M., Niemcyzk, W., Sternemann, U., Bevilacqua, M., & Emrich, H. M. (1998). Impaired binocular depth inversion in patients with alcohol withdrawal. *Journal of Psychiatric Research*, *30*(6), 469–474.

Schneider, U., Leweke, F. M., Sternemann, U., Weber, M. M., & Emrich, H. M. (1996). Visual 3D illusion: A systems-theoretical approach to psychosis. *European Archives of Psychiatry and Clinical Neuroscience, 246*, 256–260.

Semple, D. M., Ramsden, F., & McIntosh, A. M. (2003). Reduced binocular depth inversion in regular cannabis users. *Pharmacology, Biochemistry, and Behavior*, *75*(4), 789–793.

Sherman, A., Papathomas, T. V., Jain, A., & Keane, B. P. (2011). The roles of perspective, angle polarity, stereo and motion parallax in perceiving 3D objects. *Seeing and Perceiving, 25*, 263–285.

Sternemann, U., Schneider, U., Leweke, F. M., Bevilacqua, C. M., Dietrich, D. E., & Emrich, H. M. (1997). [Pro-psychotic change of binocular depth inversion by sleep deprivation]. *Nervenarzt*, *68*(7), 593–596. (in German)

Tsuruhara, A., Nakato, E., Otsuka, Y., Kanazawa, S., Yamaguchi M. K., & Hill, H. (2011). The hollow-face illusion in infancy: Do infants see a screen based rotating hollow mask as hollow? *i-Perception*, *2*, 418–427.

Wade, N. J., Campbell, R. N., Ross, H. E., & Lingelbach, B. (2010). Necker in Scotch perspective. *Perception*, *39*, 1–4.

Wheatstone, C. (1838). On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London*, *128*, 371–394.

Yellott, J. I. Jr. (1981). Binocular depth inversion. *Scientific American, 245*, 118–125.

Chapter 88
The Illusion of Sex

Richard Russell

In the Illusion of Sex (Fig. V.88-1), two faces appear side by side. One appears female, and the other appears male. In fact, both images are of the same face that has been manipulated only very slightly (this illusion was originally published in Russell, 2009). The original face is not of an actual person; it is an androgynous face that was produced by morphing together average male and average female faces. These average faces were created by morphing together 22 male faces and 22 female faces. This androgynous face is shown in Figure V.88-2. In the first version of the Illusion of Sex, the androgynous face is manipulated by making every part of the image darker (as on the right of Fig. V.88-1) or lighter (as on the left of Fig. V.88-1), except for the eyes and lips, which are left unchanged and hence are exactly the same in the two images. Although the eyes and lips appear darker in the left image than in the right image, they are not. This is an example of simultaneous contrast.

A variant of the Illusion of Sex appears in Figure V.88-3. Again, two faces appear side by side, one apparently male and the other apparently female. As in the first version of the Illusion of Sex, both faces were created by manipulating the androgynous face shown in Figure V.88-2. To create the variant seen in Figure V.88-3, the eyes and lips of the androgynous face were made lighter (as on the right of Fig. V.88-3) or darker (as on the left of Fig. V.88-3) while the rest of the image remains unchanged and hence exactly the same. The effect of this manipulation on the apparent gender of the faces in Figure V.88-3 is about the same or perhaps slightly weaker than of the manipulation shown in Figure V.88-1.

In Figures V.88-1 and V.88-3, either the eyes and lips are manipulated but not the rest of the image (as in Fig. V.88-3), or the eyes and lips are held constant while the rest of the image is manipulated (as in Fig. V.88-1). What happens if the *entire* image is made darker or lighter? Figure V.88-4 shows the androgynous face manipulated by lightening the entire image (as on the left) or darkening the entire image (as on the right). Here both images appear androgynous, or, in any event, neither image appears more male or female than the other. Manipulating the overall lightness of the face has no effect on the apparent sex of the face. We can



Figure V.88-1. The Illusion of Sex. (Figure by Richard Russell.)



Figure V.88-2. Androgynous face used to create the Illusion of Sex. (Figure by Richard Russell.)

conclude from this image that it is not darkness or lightness per se that makes the face appear to be of one sex or the other. This is also consistent with the fact that in Figure V.88-1, the face that has been darkened looks male, while in Figure V.88-2 the face that has been darkened (in the eyes and lips) looks female. Rather, it seems that the contrast between the eyes and lips and the surrounding face is critical for the illusion. The androgynous face looks male when this contrast is reduced (as on the right of Figs. V.88-1 and V.88-3) and it looks female when this contrast is increased (as on the left of Figs. V.88-1 and V.88-3).

The Illusion of Sex works by manipulating the luminance contrast between the eyes and lips and the rest of the face. I have named this kind of contrast "facial contrast" and shown that manipulating it affects the attractiveness of male and female faces differently (Russell, 2003). In particular, female faces are more attractive with facial contrast increased than decreased, while male faces are more attractive with facial contrast decreased than increased. This effect of facial contrast on attractiveness is a result of a naturally occurring sex difference in facial contrast. In Caucasian and East Asian faces, females have greater facial contrast than males (Russell, 2009).

This physical sex difference in facial contrast has perceptual consequences. Humans use facial contrast as a cue for classifying faces by sex (Dupuis-Roy, Fortin, Fiset, & Gosselin, 2009; Nestor & Tarr, 2008). Facial contrast also affects judgments of the masculinity or feminity of actual faces. Female faces with higher facial contrast are rated as more feminine than those with lower facial contrast, while male faces with lower facial contrast are rated as more masculine than those with higher facial contrast (Russell, 2009). The Illusion of Sex, by manipulating facial contrast, demonstrates that there is a causal relationship between facial contrast and the apparent gender of a face. Thus there is a naturally occurring physical difference between male and female faces in their amount of facial contrast, and this difference is used by the human visual system in perceiving the sex of other human faces. This is consistent with the notion that the visual system is sensitive to statistical regularities in the visual environment and uses them for perceiving the world, in this case for perceiving whether a person is male or female.

The sex difference in facial contrast exists because female skin is lighter than male skin (reflecting 2% to 3% more light), a situation found in all racial groups around the world (for reviews, see Frost, 1988; Jablonski & Chaplin, 2000). Though female skin is lighter than male skin, female eyes and lips are not lighter than male eyes and lips, or they are only slightly lighter. To date, facial contrast has been measured only in Caucasian and East



Figure V.88-3. Alternate version of the Illusion of Sex. (Figure by Richard Russell.)



Figure V.88-4. Androgynous face made lighter or darker, without altering facial contrast. (Figure by Richard Russell.)

Asian faces (Russell, 2009). Eyes and lips are darker than the surrounding skin (Sinha, 2002; only Caucasian, East Asian, and South Asian faces were tested), which is the underlying reason why facial contrast exists. If the skin is lighter but the eyes and lips are not, it means there will be greater facial contrast. This is the situation with female faces relative to male faces. This means that the first version of the Illusion of Sex (Fig. V.88-1) is in some sense not an illusion at all—we are simply making the face more male-typical by making the skin darker or more female-typical by making the skin lighter, all the while keeping the eyes and lips about the same. There are many features that differ between male and female faces, such as the protrusion of the brow, the length of the lower face, and the height and thickness of the eyebrows (Burton, Bruce, & Dench, 1993). I predict that manipulating any of these sexually dimorphic features in an androgynous face would have similar effects (though perhaps different effect sizes) as the manipulation of facial contrast in the Illusion of Sex.

The second version of the illusion (Fig. V.88-3) is more clearly an illusion in that what is being changed (the luminance of the eyes and lips) does not actually differ between male and female faces. This version of the illusion works because the manipulation still affects facial contrast. In this way, the Illusion of Sex demonstrates the general reliance of the visual system on contrast rather than absolute luminance and the reliance on contrast in the specific context of face perception.

The Illusion of Sex is relevant not only for understanding visual sex classification but also for understanding the use of cosmetics. Mascara, eye liner, and eye shadow all typically darken the eyes. It is exceptionally rare for these products to come in shades that actually make the eyes lighter rather than darker (hence the term eye *shadow*). Lipstick is used to darken the lips as well as to make them redder. It is very unusual for a shade of lipstick to make the lips appear lighter than they naturally are. However, the luminance of the skin is typically not darkened by cosmetics, and in some periods and cultures the skin has been lightened through cosmetics. Thus makeup is typically worn in ways that increase facial contrast by making the eyes and lips darker while keeping the luminance of the skin constant or even making it lighter. In this way, typical makeup use changes the face in the same way that the androgynous face is changed to make the left images in Figure V.88-3. By increasing facial contrast, cosmetics are used in precisely the correct way to make the pigmentation of the face appear more feminine. Indeed, it has been shown that the same face wearing cosmetics has greater facial contrast than it does without cosmetics, whether the makeup was applied by the wearer (Russell 2009) or by a professional cosmetic artist (Etcoff, Stock, Haley, Vickery, & House, 2011).

There are many ways that cosmetics *could* be applied (nosestick, chin shadow, etc.) but are not. It is very unlikely that the consistent increase in facial contrast achieved through cosmetics use is an accident. And there is a very plausible explanation *why* cosmetics are used in precisely the correct way as to increase facial contrast. As the Illusion of Sex demonstrates, increasing facial contrast makes a face appear more female and more feminine. Femininity plays a large role in female facial attractiveness (reviewed by Rhodes 2006), with more feminine faces appearing more attractive. Thus increasing facial contrast should increase the femininity of a face and hence the attractiveness. Consistent with this notion, several studies have shown that cosmetics increase the attractiveness and femininity of a face (reviewed by Russell, 2010). In a recent study, different amounts of makeup were applied to each of a set of target faces by a cosmetic artist, and facial contrast was found to be increased more by heavy application of makeup than light application of makeup (Etcoff et al., 2011). These faces were rated for attractiveness, and it was confirmed that facial contrast mediates the link between cosmetics use and attractiveness.

The Illusion of Sex highlights a statistical regularity that is fundamental to the perception of gender—the sex difference in facial contrast. Further, it demonstrates how visual illusions can be useful not only for increasing our understanding of the visual system but also for shedding light on fundamental human behaviors such as the use of cosmetics.

REFERENCES

Burton, A. M., Bruce, V., & Dench, N. (1993). What's the difference between men and women? Evidence from facial measurement. *Perception, 22*, 153–176.

Dupuis-Roy, N., Fortin, I., Fiset, D., & Gosselin, F. (2009). Uncovering gender discrimination cues in a realistic setting. *Journal of Vision*, *9*(2), 10.

Etcoff, N. L., Stock, S., Haley, L. E., Vickery, S. A., & House, D. M. (2011). Cosmetics as a feature of the extended human phenotype: Modulation of the perception of biologically important facial signals. *PLoS ONE*, *6*(10), e25656.

Frost, P. (1988). Human skin color: A possible relationship between its sexual dimorphism and its social perception. *Perspectives in Biology and Medicine*, *32*, 38–58.

Jablonski, N. G., & Chaplin, G. (2000). The evolution of human skin coloration. *Journal of Human Evolution*, *39*, 57–106.

Nestor, A., & Tarr, M. J. (2008). Gender recognition of human faces using color. *Psychological Science*, *19*(12), 1242–1246.

Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology, 57*, 199–226.

Russell, R. (2003). Sex, beauty, and the relative luminance of facial features. *Perception*, *32*, 1093–1107.

Russell, R. (2009). A sex difference in facial contrast and its exaggeration by cosmetics. *Perception*, *38*, 1211–1219.

Russell, R. (2010). Why cosmetics work. In R. Adams, N. Ambady, K. Nakayama, & S. Shimojo (Eds.), *The science of social vision*. New York: Oxford University Press. pp. 186–203

Sinha, P. (2002). Qualitative representations for recognition. *Lecture Notes in Computer Science*, *2525*, 249–262.

Chapter 89
The Bogart Effect

Sharon Gilad-Gutnick, Rohan Varma, and Pawan Sinha

A key aspect of human face perception is our ability to accurately estimate the direction of a gaze (Hinde & Rowell, 1962; Kleinke, 1986). Studies have shown that besides using the overall orientation of the head, the visual system relies on cues from the eyes (Baron-Cohen, 1995; Perrett et al. 1985; Todorović, 2006; von Grunau & Anston, 1995). Even under highly degraded conditions, the eyes seem to provide a wealth of information for determining gaze direction. Cline (1967) showed that observers can detect gaze deviations as small as 1.4° at a distance of a meter. This ability is not surprising given the key role that gaze direction plays in mediating social interactions and enabling joint attention (e.g., see Driver, Davis, & Ricciardelli, 1999; Emery, 2000; Hietanen, 1999, for review). For instance, direct gaze, even from an image, can be very persuasive (e.g., posters of Uncle Sam calling people to join the army during World War II).

So how do we decide where someone is looking? Many researchers have suggested that the unique morphology of the human eye means that useful information can be recovered using simple structure-based cues (Langton, 2000, p. 52). That is, compared to other animals, the human eye has a relatively small dark region (including the pupil and iris) and a large white sclera on either side of the iris (Kobayashi & Kohshima, 1997). The high contrast between the white sclera and the dark iris make the discrimination of gaze direction much easier in humans than in other animals and implies that simple geometric features can be used to determine eye-gaze direction. One geometry-based strategy that many researchers have suggested plays a predominant role in the determination of gaze direction is iris eccentricity, or the relative position of the circular iris within the visible part of the sclera (Anstis, Mayhew, & Morley, 1969, p. 489; Baron-Cohen, 1994, p. 519; Symons, Lee, Cedrone, & Nishimura, 2004, p. 452). According to this idea, the position and curvature direction of the high contrast between the sclera and the iris facilitates the extraction of geometric cues considered essential for estimating gaze direction.

While a geometry-based eye-gaze estimation strategy has been the basis of many theories, such a strategy would rely on relatively detailed curvature information and would therefore function suboptimally under low-resolution viewing conditions. Partly in response to this concern, the past decade has seen the rise of luminance-based theories of eye-gaze estimation. The idea of luminance-based estimation of gaze direction arose from the observation that contrast negation affects eye-gaze perception, and an early demonstration and possible explanation for this phenomenon was offered by Sinha (2000) and named the "Bogart effect."

The Bogart effect received its name after being first demonstrated using an image of Humphrey Bogart (Sinha, 2000). In the untouched version of the image, Humphrey Bogart is seen looking to his left. However, when the same image is contrast negated (like in negatives of a photographic film), Bogart's gaze appears to reverse direction and look to the right (Fig. V.89-1). As can be seen from Figure V.89-1a, the effect of contrast negation on the estimation of gaze direction is not unique to Bogart's face.

To test the robustness of this effect, subjects were shown the images from Figure V.89-1 and asked to indicate whether the individuals depicted in the image appeared to be looking to the (subject's) left or right. As Figure V.89-2 shows, contrast negation of the entire face led to a reversal of the perceived gaze direction in most subjects. For the few cases in which contrast negation did not lead to a reversal in the estimation of gaze direction, postexperiment debriefing suggested that subjects might have been deliberately compensating for the contrast negation and estimating the "real" gaze direction by reversing the perceived one. Overall, the strong correlation of perceived gaze direction with contrast polarity in these images suggests that the visual system uses a gaze-estimation strategy that may be based on a simple luminance-based heuristic: Which side is darker?

The perceptual power of this rule of thumb is evident in the observation that, for most subjects, gaze-direction reversal was obtained even though the subjects could tell that they were looking at contrast-negative images and thus "knew" that the irises were the light regions in the eyes. That is, the gaze-estimation process appears to be largely "cognitively impenetrable" (Pylyshyn, 1999), with the rule of thumb taking precedence over cognitive classifications of eye regions as the iris and sclera. Importantly, gaze-direction reversal was obtained despite the fact that the geometric cues obtained from the negated image did not coincide with the new perceived direction. That is, the putative iris in the negated image is of an abnormal size



Figure V.89-1. The influence of contrast negation on perceived gaze direction. (Adapted from Sinha, 2000.)

relative to the rest of the eye and/or is bounded by arcs that curve in the wrong direction. Thus it seems that an eye region can be treated as an iris by the gaze-estimation process, so long as it satisfies the darkness heuristic, even though its other structural attributes may be abnormal. In other words, the Bogart effect suggests that rather than using the position or curvature of high contrast between the sclera and the iris to judge gaze direction, the visual system may use the polarity of high contrast to accurately perform this task.

A luminance-based heuristic, as proposed by Sinha (2000), provides the missing link for similar findings regarding the effect of contrast negation on the estimation of gaze direction. For instance, Ricciardelli, Baylis, and Driver (2000) showed that subjects were very poor at correctly judging the gaze direction of images with contrast-reversed eyes. They noted that a system based only on the position of the contrast boundary cannot explain why the reversal of contrast around the eyes should affect gaze perception and proposed that their finding can be explained by an "expert" system, whereby the darker region of the eye is the part that does the looking. The Bogart effect, in essence, demonstrates the validity of this heuristic for highly eccentric gaze directions by showing that contrast negation actually reverses the direction of the estimated gaze by this "expert" system.

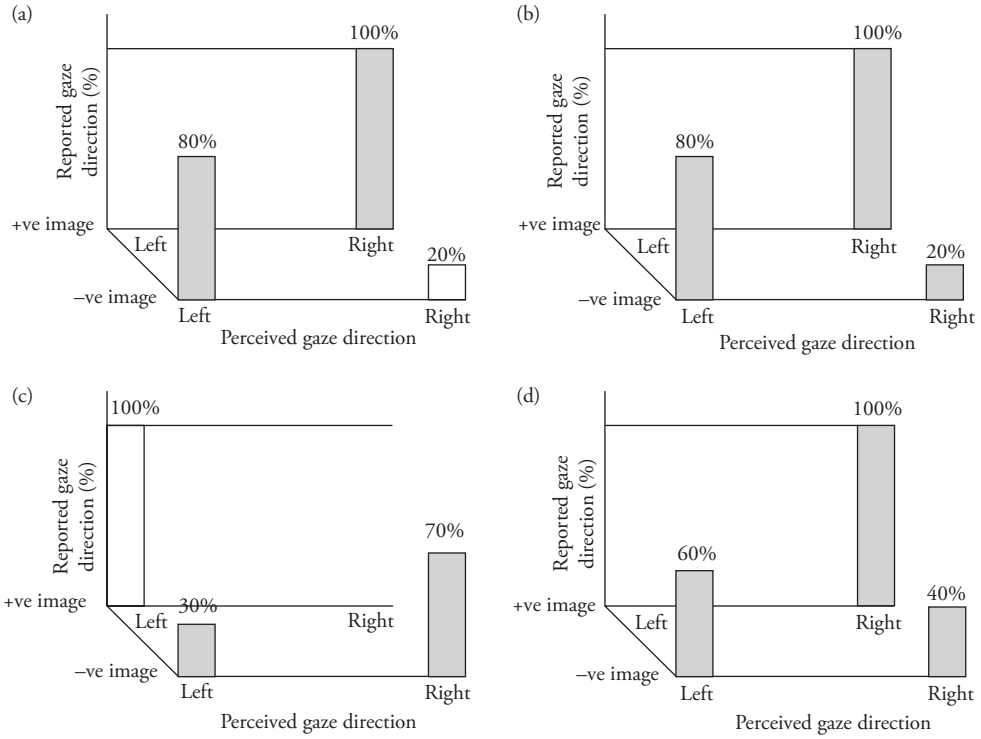


Figure V.89-2. Gaze-estimation results from 10 subjects tested with the face images shown in Figure V.89-1. Panels (a) through (d) display results for the positive and negative images of Figures V.89-1a through V.89-1d, respectively. Contrast-negation reverses perceived gaze direction for the majority of the subjects. The gaze direction reversal effect is weaker in (d), perhaps because the gaze direction is close to frontal. (Adapted from Sinha, 2000.)

The use of this simple polarity-based strategy may also account for the bloodshot illusion, first demonstrated by Ando and Osaka (1998). Gaze direction of forward-looking eyes seems to be skewed to the side if one scleral region is darkened (as if bloodshot). One possible explanation for this phenomenon is that the visual system is grouping both the iris and the darkened scleral region together into one "composite" iris (since both of them now satisfy the darkness heuristic). The inclusion of the "bloodshot" sclera makes the center of this composite iris shift to a side, thus leading to a change in the perceived gaze direction (see also Ando, 2002). Data from a follow-up study led Ando (2004) to propose a similar luminance-based based strategy based on the ratio between the sclera and its surrounding region rather than the polarity within the eye. Rather than darkening the sclera within the eye, Ando artificially placed a dark square in the area surrounding one corner of the eye. He found that eye gaze seemed to be directed toward the side for which the luminance ratio between the sclera and its surrounding skin region was smaller. One important advantage of the strategy proposed by Ando is that it promotes correct estimates of gaze direction under different lighting conditions (e.g., a shadow cast on one side of the eye would not cause errors in perceived gaze direction).

In the discussion so far we have emphasized a luminance-based approach to estimating eye gaze. However, it is clear that the role of geometry-based cues cannot be completely discounted. After all, an eye represented only by the outlines of an oval and a circle, with no luminance contrast across the iris and scleral regions, can still provide sufficient cues to determine gaze direction. It is possible that luminance and geometry are both important and are combined in ways that are yet to be fully understood to enhance the accuracy of gaze estimates. Indications of such cue interactions come from studies that have used a cue-conflict paradigm. For instance, measuring how subjects judge the direction of eye gaze shown at different eccentricities in positive and negative polarity images, Olk, Symons, and Kingstone (2008) found that the highest number of errors were incorrect "straight" judgments at low eccentricities, suggesting that gaze judgments were based on the outcome of a competition between the luminance cues that indicate that the gaze is directed from the dark part of the eye and geometric cues that indicate the opposite. Indeed, they found a reduced error rate and fewer incorrect "straight" responses for more extreme gaze directions. In addition to the use of geometric- and/or luminance-based information from the eyes, it is clear that the orientation of the head also influences the perception of general gaze direction (Todorović, 2006). However, in the absence of head-orientation cues (such as nearly frontal gaze), further experiments are needed in order to fully resolve the conditions under which a competition between different cues from the eyes might ensue, as well as how the final decision for which strategy to use is ultimately made.

Although both geometric and luminance cues are likely to contribute to gaze estimates, the latter probably dominate over long viewing distances. An informal study we conducted illustrates this point. Participants were shown line drawings of faces that fell into two categories: those with a darkened iris (providing cues for "which side is darker?") and those with an empty iris (providing only structural

cues). Gaze directions were approximately 30° to the left or right. Starting from a distance at which gaze could not be estimated in either of the two stimulus conditions (filled or unfilled irises), 10 participants were instructed to walk slowly toward the face and indicate when they were able to determine the direction of the eye gaze. This very simple procedure yielded a clear difference between the threshold distance for the two conditions: with filled irises, observers were able to reliably estimate gaze direction from a viewing distance 50% longer than that corresponding to the unfilled iris condition. Thus luminance-based cues likely play a key role in gaze estimation under conditions in which the contours of the eyes are blurred, such as when viewing a face from a distance. By not being overly reliant on structural cues alone, the visual system is able to estimate gaze direction even under degraded viewing conditions caused by refractive errors or large viewing distances. The rule of thumb for gaze estimation based on ordinal brightness relationships ("the darker side is gazing") allows for a simple and robust implementation of this important perceptual ability.

REFERENCES

Ando, S. (2002). Luminance-induced shift in the apparent direction of gaze. *Perception, 31*(6), 657–674.

Ando, S. (2004). Perception of gaze direction based on luminance ratio. *Perception, 33*(10), 1173–1184.

Ando, S., & Osaka, N. (1998). Bloodshot illusion: Luminance affects perceived gaze-direction. *Investigative Ophthalmology & Visual Science, 39*(4), 172.

Anstis, S. M., Mayhew, J. W., & Morley, T. (1969). The perception of where a face or television "portrait" is looking. *The American Journal of Psychology, 82*(4), 474–489.

Baron-Cohen, S. (1994). How to build a baby that can read minds: Cognitive mechanisms in mindreading. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition, 13*(5), 513–552.

Baron-Cohen, S. (1995). The eye direction detector (EDD) and the shared attention mechanism (SAM): Two cases for evolutionary psychology. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its role in origins and development* (pp. 41–59). Hillsdale, NJ: Lawrence Erlbaum.

Cline, M. G. (1967). The perception of where a person is looking. *The American Journal of Psychology, 80*(1), 41–50.

Driver, J., Davis, G., & Ricciardelli, P. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition, 6*(5), 509–540.

Emery, N. J. (2000). The eyes have it: The neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews, 24*(6), 581–604.

Hietanen, J. K. (1999). Does your gaze direction and head orientation shift my visual attention? *NeuroReport, 10*(16), 3443–3447.

Hinde, R. A., & Rowell, T. E. (1962). Communication by postures and facial expressions in the rhesus monkey (Macaca mulatta). *Proceedings of the Zoological Society of London, 138*(1), 1–21.

Kleinke, C. L. (1986). Gaze and eye contact: A research review. *Psychological Bulletin, 100*(1), 78–100.

Kobayashi, H., & Kohshima, S. (1997). Unique morphology of the human eye. *Nature, 387*(6635), 767–768.

Langton, S. R. H. (2000). The mutual influence of gaze and head orientation in the analysis of social attention direction. *The Quarterly Journal of Experimental Psychology Section A, 53*(3), 825–845.

Olk, B., Symons, L., & Kingstone, A. (2008). Take a look at the bright side: Effects of contrast polarity on gaze direction judgments. *Attention, Perception, & Psychophysics, 70*(7), 1298–1304.

Perrett, D. I., Smith, P. A. J., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D. & Jeeves, M. A. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society of London, B, 223*(1232), 293–317.

Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences, 22*(3), 341–365.

Ricciardelli, P., Baylis, G., & Driver, J. (2000). The positive and negative of human expertise in gaze perception. *Cognition, 77*(1), B1–B14.

Sinha, P. (2000). Last but not least: Here's looking at you, kid. *Perception, 29*(8), 1005–1008.

Symons, L. A., Lee, L., Cedrone, C. C., & Nishimura, M. (2004). What are you looking at? Acuity for triadic eye gaze. *Journal of General Psychology, 131*(4), 451–469.

Todorović, D. (2006). Geometrical basis of perception of gaze direction. *Vision Research, 46*(21), 3549–3562.

von Grunau, M., & Anston, C. (1995). The detection of gaze direction: A stare-in-the-crowd effect. *Perception, 24*(11), 1297–1313.

Chapter 90
The Presidential Illusion

Sharon Gilad-Gutnick and Pawan Sinha

Figure V.90-1 shows images of 15 women. It is not hard to accept that these are distinct women, different along many dimensions. Remarkably, however, all of these images in fact depict the same woman wearing different disguises. By changing makeup and external features such as hair and accessories, this master of disguise does more than reshape her image; she somehow seems to invent a whole new person. The details of the eyes and nose, and their mutual spatial relationships, stay largely the same across these images, and yet "superficial" changes have the ability to suggest new identities. The effectiveness of these disguises provides a hint that external head features might be quite important for conveying facial identity.

In the past, researchers of human vision as well as developers of computer vision have focused primarily on the shape and spatial arrangement of internal facial measurements (Biederman, 1987; Bruce, Doyle, Dench, & Burton, 1991; Carey & Diamond, 1977; Haig, 1984; Hoffman & Richards, 1984). External features, such as the overall shape of the face, haircut, and jawline, are typically excluded (Goffaux & Rossion, 2007; Leopold, Rhodes, Müller, & Jeffery, 2005; Robbins & McKone, 2003), covered (Duchaine & Nakayama, 2006; Le Grand, Mondloch, Maurer, & Brent, 2004; Mondloch, Le Grand, & Maurer, 2002), or controlled by combining the internal features of different faces with the same external features (Yovel & Kanwisher, 2004, 2005). Such methodologies have become commonplace because it is widely believed that external features change dramatically when people change their hair or wear a hat, and therefore external features are not likely be an essential part of the face-recognition mechanism.

In addition to being consistent with our perceptual intuition, the designation of the eyes, nose, and mouth as critical facial features is the outcome of computational studies of face representation (Bartlett, Movellan, & Sejnowski, 2002; Lee & Seung, 1999). In most computer vision systems, these features include distances and angles between—and local intensity patches around—the centroids of the major nameable internal features (Campbell, Cannon, Jones, & Morgan, 1987; Kanade, 1973). The use of external features has been deliberately omitted from most automatic face-recognition algorithms because they are considered too variable. Also, external regions are more difficult to align, and their reliable extraction is itself a major challenge. As a result, machine-based face-recognition systems are typically designed to rely primarily, even solely, on internal cues. However, these systems do not perform very robustly in practice, as is evident from the minimal deployment of such systems in real-world settings such as airports. Clearly some critical aspects of facial identity are missing from these computational classification systems. Recent studies of the human visual system may help us fill in the missing pieces of this puzzle.

Most readers who look at the left panel of Figure V.90-2 will see ex-president Bill Clinton and vice president Al Gore. However, look again. Both men actually have identical internal features—those of President Clinton; in fact, the internal features of the vice president have been entirely replaced with those of the president. The fact that most people fail to notice this manipulation, even after being told of it, highlights the importance of external facial cues for face recognition. This illustration, named the 'Presidential Illusion', brought to the table the idea that, unlike computer vision systems and many models of human face processing, the human visual system actually seems to analyze the structure of the entire head. Now look at the right panel of Figure V.90-2. Despite having just seen the same demonstration, the reader may not notice that, once again, the vice president's internal features have been replaced with those of the president (Sinha & Poggio, 2002). This repeated demonstration of the presidential illusion provides evidence that the original Clinton–Gore illusion is not one of a kind but is rather a robust phenomenon that reflects a general principle regarding the role of head shape in modifying internal cues for facial encoding.

The effectiveness of the presidential illusion underscores the important point that by excluding external facial features, such as the head and hair shape, we lose critical information about the way faces are represented in real life. An interesting question to consider, therefore, is whether whole-head processing is a general principle that can be extended to all face processes or if it specifically reflects the nature of facial encoding used by the visual system for the identification of individuals. For example, would supplementing the internal features of one face with those of another affect the perception of other common facial attributes, such as gender, race, or age?



Figure V.90-1. Fifteen different faces that, in fact, are all the same woman in various disguises. (Images created by Nienke Klunder were taken from the Chive.com and used under the US Copyright Fair Use Act.)

Figure V.90-3 shows the presidential illusion adapted for gender; the woman in the picture has been given the man's internal features. Embedding the internal features in the external features of the opposite gender seems to reverse perceived gender of those internal features.

Consistent with the presidential illusion, a growing body of perceptual evidence has indicated that external features play a significant role in face recognition by the human visual system (Ellis, Shepherd, & Davies, 1979; Haig, 1986; Shepherd, 1981). Studies on the relative roles of internal and external cues have given rise to a "feature hierarchy," or the idea that certain features are more important to face recognition than others (Fraser, Craig, & Parker, 1990). One such feature hierarchy study found the head outline to be the most significant cue, followed by the eyes, the mouth, and then the nose, resulting in the external features being named "cardinal" (Ellis, 1986). Additionally, recent reports have shown that the fusiform face area, a brain region primarily studied for its role in representing internal facial features, also encodes information about external cues (Andrews, Davies-Thompson, Kingstone, & Young, 2010; Betts & Wilson, 2010). Notably, researchers have repeatedly found that while internal features are more useful than external features for recognizing a familiar face, recognizing an unfamiliar face relies extensively on external features (Bruce et al., 1999; Ellis et al., 1979; Shepherd, 1981; Young, Hay, McWeeny, Flude, & Ellis, 1985). An interesting aspect of this finding is that, even for familiar faces, the relative importance of these two feature sets becomes reversed as resolution decreases (e.g., at viewing distances larger than 20 feet; Jarudi & Sinha, 2003). The idea that face recognition at low resolutions relies heavily on external cues is well in line with developmental studies showing that face discrimination by infants, whose visual acuity is limited, depends more on external features than on internal ones (Bushneil, Sai, & Mullin, 1989; Field, Cohen, Garcia, & Greenberg, 1984; Pascalis, de Schonen, Morton, Deruelle, & Fabre-Grenet, 1995).



Figure V.90-2. The presidential illusion. In both panels, the vice presidents have been given the presidents' internal facial features in their original spatial configuration. (Left: From Sinha & Poggio, 1996; Right: After Sinha & Poggio, 2002.)

However, the contribution of external cues to face recognition is not found only in degraded conditions or for an unfamiliar face. In fact, it seems that external features, while poor indicators of identity on their own, can provide an important frame of reference for analyzing facial identity. An imaging study by Axelrod and Yovel (2010) showed that the fusiform face area is not only sensitive to external cues but is also sensitive to how it influences the identity of internal features. Their finding is consistent with a previous behavioral study showing that subjects make more recognition errors when the same face is presented with different hairstyles (Patterson & Baddeley, 1977). Finally, a study by Jarudi and Sinha (2003) showed that for distant faces it is neither the internal nor the external configurations on their own that subserve recognition but rather measurements corresponding to how internal features are placed relative to external features. This idea is consistent with the well-established notion of holistic face processing, according to which the brain represents faces not as decomposed parts but rather as a collective set of features (Farah, Wilson. Drain, & Tanaka, 1998; Le Grand et al., 2004; Tanaka & Farah, 1993; Young, Hellawell, & Hay, 1987).

Although a strategy that relies strictly on external cues would be quite limited, the integration of external cues into a face-identification strategy is advantageous for correct recognition under a variety of viewing conditions. One such benefit is the saliency of the external features for use with reduced resolution, such as for recognition from large distances. However, under certain conditions, the integration of the external cues into an overall head representation could actually interfere with correct face identification. Consider our example in Figure V.90-1—even though we can, with



Figure V.90-3. External features can override internal cues for gender identification. The woman's internal features have been replaced with the man's internal features (from above the eyebrows to below the bottom lip). (Original image from Hulton Archive/Getty Images.)



Figure V.90-4. Robustness of human face perception. Subjects were able to recognize more than half of all familiar faces shown to them at the resolution depicted here. Individuals shown, in order, are: Michael Jordan, Woody Allen, Goldie Hawn, Bill Clinton, Tom Hanks, Saddam Hussein, Elvis Presley, Jay Leno, Dustin Hoffman, Prince Charles, Cher, and Richard Nixon. (From Sinha et al., 2006.)

deliberate effort, convince ourselves that all of the images show the same woman, it seems that for a casual observer, her external changes are effective in conveying different identities. It is possible that the visual system resolves this paradox via a competition of strategies or a compensatory process: it may initially use a default strategy by which it utilizes integrated information from both internal and external cues, but when unsuccessful, it is able to resort to an alternative strategy and "ignore" the external cues to resolve the true identity based only on the internal features. There may also be conditions for which the brain initially chooses to use one strategy over another; for example, it has been shown that the ability to tolerate degradation increases with familiarity (Sinha, Balas, Ostrovsky, & Russell, 2006), suggesting a flexible strategy that preferentially utilizes more internal features for high-quality faces and external features for low-resolution faces.

The compelling role of external features for recognition under realistic viewing conditions has motivated an interest in their use for automatic face classification. Correctly interpreting degraded images is especially important in automated vision systems, since in many cases these systems acquire images at low resolutions due to hardware limitations and large viewing distances. Most facial classification algorithms still perform suboptimally compared to humans because they rely primarily on part-based strategies (Brunelli & Poggio, 1993; Doi, Sato, & Chihara, 1998), which are especially limited in their ability to handle degraded viewing conditions that reduce feature detail. For instance, consider Figure V.90-4. Even though human observers perform well at recognizing such impoverished images (Bachmann, 1991; Bruce et al., 1999; Harmon & Julesz, 1973; Sinha & Poggio, 2002), performance of even the best existing machine vision system falls dramatically short. In this case, an approach for interpreting low-resolution images might be derived from the human visual system. Indeed, a few automatic classification algorithms have already been designed based on a "holistic" strategy (e.g., the Eigenface algorithm; Turk & Pentland, 1991).

So might a machine vision system soon identify faces holistically and as well as a human? Several hurdles still exist. Foremost, it is unclear if modeling an automated system after human vision is the most effective scheme, because the two systems may have a different set of functional requirements and limitations. For instance, a holistic-based method must rely on matching templates, resulting in a computationally expensive algorithm that would actually be less robust to position and scale invariance than a feature-based strategy (for a review, see Tsao & Livingstone, 2008). Additionally, in the face-classification field there are many security-related applications in which the reliability obtained by internal cues is essential, since external cues are more variable and easy to disguise. Most important, however, even within the field of human vision research, we have yet to provide a comprehensive description of the nature of holistic face representations.

Overall, the presidential illusion, as well as many subsequent findings, brought to the forefront of vision research the idea that a face is often perceived as a whole entity, whereby information from both internal and external facial cues are integrated. However, many questions still remain unanswered: Precisely how do external features modify the identity of internal features? How can this perceptual idea be translated into computational terms? Finally, when a future automated system is developed that is so human-like that it too succumbs to the presidential illusion, would we consider this a success or failure?

REFERENCES

Andrews, T. J., Davies-Thompson, J., Kingstone, A., & Young, A. W. (2010). Internal and external features of the face are represented holistically in face-selective regions of visual cortex. *The Journal of Neuroscience, 30*(9), 3544–3552.

Axelrod, V., & Yovel, G. (2010). External facial features modify the representation of internal facial features in the fusiform face area. *NeuroImage, 52*(2), 720–725.

Bachmann, T. (1991). Identification of spatially quantised tachistoscopic images of faces: How many pixels does it take to carry identity? *European Journal of Cognitive Psychology, 3*(1), 87–103.

Bartlett, M. S., Movellan, J. R., & Sejnowski, T. J. (2002). Face recognition by independent component analysis. *IEEE Transactions on Neural Networks and Learning Systems, 13*(6), 1450–1464.

Betts, L. R., & Wilson, H. R. (2010). Heterogeneous structure in face-selective human occipito-temporal cortex. *Journal of Cognitive Neuroscience, 22*(10), 2276–2288.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*(2), 115–147.

Bruce, V., Doyle, T., Dench, N., & Burton, M. (1991). Remembering facial configurations. *Cognition, 38*(2), 109–144.

Bruce, V., Henderson, Z., Greenwood, K., Hancock, P. J. B., Burton, A. M., & Miller, P. (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied, 5*(4), 339–360.

Brunelli, R., & Poggio, T. (1993). Face recognition: Features versus templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *15*(10), 1042–1052.

Bushneil, I. W. R., Sai, F., & Mullin, J. T. (1989). Neonatal recognition of the mother's face. *British Journal of Developmental Psychology, 7*(1), 3–15.

Campbell, R. A., Cannon, S., Jones, G., & Morgan, N. (1987). Individual face classification by computer vision. In *Proceedings Conference on Modeling Simulation Microcomputer*, 62–63.

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. *Science, 195*(4275), 312–314.

Doi, M., Sato, K., & Chihara, K. (1998). A robust face identification against lighting fluctuation for lock control. In *Proceedings: Third IEEE International Conference on Automatic Face and Gesture Recognition* (pp. 42–47). Los Alamitos, CA: IEEE Computer Society Press.

Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia, 44*(4), 576–585.

Ellis, H. D. (1986). Face recall: A psychological perspective. *Human Learning: Journal of Practical Research & Applications, 5*(4), 189–196.

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. *Perception, 8*(4), 431–439.

Farah, M. J., Wilson. K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review, 105*(3), 482–498.

Field, T. M., Cohen, D., Garcia, R., & Greenberg, R. (1984), Mother-stranger face discrimination by the newborn. *Infant Behavior and Development, 7*(1), 19–25.

Fraser, I. H., Craig, G. L., & Parker, D. M. (1990). Reaction time measures of feature saliency in schematic faces. *Perception, 19*(5), 661–673.

Goffaux, V., & Rossion, B. (2007). Face inversion disproportionately impairs the perception of vertical but not horizontal relations between features. *Journal of Experimental Psychology: Human Perception and Performance, 33*(4), 995–1002.

Haig, N. D. (1984). The effect of feature displacement on face recognition. *Perception, 13*(5), 505–512.

Haig, N. D. (1986). Exploring recognition with interchanged facial features. *Perception, 15*(3), 235–247.

Harmon, L. D., & Julesz, B. (1973). Masking in visual recognition: Effects of two-dimensional filtered noise. *Science, 180*(4091), 1194–1197.

Hoffman, D. D., & Richards, W. A. (1984). Parts of recognition. *Cognition, 18*(1–3), 65–96.

Jarudi, I. N., & Sinha, P. (2003). *Relative contributions of internal and external features to face recognition*. Fort Belvoir, VA: Defense Technical Information Center.

Kanade, T. (1973). *Computer recognition of human faces*. Basel, Switzerland: Birkhauser.

Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature, 401*(6755), 788–791.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2004). Impairment in holistic face processing following early visual deprivation. *Psychological Science, 15*(11), 762–768.

Leopold, D. A., Rhodes, G., Müller, K.-M., & Jeffery, L. (2005). The dynamics of visual adaptation to faces. *Proceedings of the Royal Society B: Biological Sciences, 272*(1566), 897–904.

Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception, 31*(5), 553–566.

Pascalis, O., de Schonen, S., Morton, J., Deruelle, C., & Fabre-Grenet, M. (1995). Mother's face recognition by neonates: A replication and an extension. *Infant Behavior and Development, 18*(1), 79–85.

Patterson, K. E., & Baddeley, A. D. (1977). When face recognition fails. *Journal of Experimental Psychology: Human Learning and Memory, 3*(4), 406–417.

Robbins, R., & McKone, E. (2003). Can holistic processing be learned for inverted faces? *Cognition, 88*(1), 79–107.

Shepherd J. W. (1981). Studies of cue saliency. In G. Davies, H. Ellis, & J. Shepherd (Eds.), *Perceiving and remembering faces* (pp. 105–131). New York: Academic Press.

Sinha, P., Balas, B., Ostrovsky, Y., & Russell, R. (2006). Face recognition by humans: Nineteen results all computer vision researchers should know about. *Proceedings of the IEEE, 94*(11), 1948–1962.

Sinha, P., & Poggio, T. (2002). United we stand. *Perception, 31*(1), 133.

Sinha, P., & Poggio, T. (1996). I think I know that face. *Nature, 384*(6608), 404.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology Section A, 46*(2), 225–245.

Tsao, D. Y., & Livingstone, M. S. (2008). Mechanisms of face perception. *Annual Review of Neuroscience, 31*, 411–437.

Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience, 3*(1), 71–86.

Young, A. W., Hay, D. C., McWeeny, K. H., Flude, B. M., & Ellis, A. W. (1985). Matching familiar and unfamiliar faces on internal and external features. *Perception, 14*(6), 737–746.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16*(6), 747–759.

Yovel, G., & Kanwisher, N. (2004). Face perception: Domain specific, not process specific. *Neuron, 44*(5), 889–898.

Yovel, G., & Kanwisher, N. (2005). The neural basis of the behavioral face-inversion effect. *Current Biology, 15*(24), 2256–2262.

Chapter 91
About Face

The Margaret Thatcher Illusion

Peter Thompson

THE ILLUSION

In May 1979 Margaret Thatcher's Conservative Party won the general election and she became the first female prime minister in the United Kingdom. A year later I was looking for pictures of a familiar face and acquired two posters of Thatcher from my local Conservative Association. My aim was to produce a demonstration for students, to show that at a distance two images might look identical while up close they might look very different. Fine details (or high spatial frequencies, as vision scientists call them) in a distant image will fall beyond the limit of our visual acuity. However, as the image comes closer, those fine details can fall within our "window of visibility." Thus my intention was to take one image of Margaret Thatcher and alter its appearance by changing some of the fine details. This I did by cutting out the eyes and mouth of the picture and rotating them by 180°. The result was a gratifyingly grotesque image. Surely viewed from afar the undoctored image and my disfigured image would look identical, but up close the two images would look horrendously different. I have never discovered whether this demonstration works because a totally unexpected result of my disfiguring Margaret Thatcher intervened, as shown in Figure V.91-1. Viewing, by chance, my two images upside-down, I noticed that they now looked very similar, while right-way-up they looked very different.

Pleased with the effect I had inadvertently created, I dashed off a brief note (242 words) to the journal *Perception*. The paper was published (Thompson, 1980), and I thought little more about it for some time. At the time of this writing, the paper had been cited some 272 times (or 381 times according to the more catholic Google Scholar). As Figure V.91-2 shows, there is little evidence that its popularity is in decline. A somewhat fuller story of the discovery of the illusion can be found in Thompson, Anstis, Rhodes, Jeffery, and Valentine (2009).

WHAT DOES THE THATCHER ILLUSION TELL US ABOUT FACE PROCESSING?

The study of upside-down faces has a long history; Giuseppe Arcimboldo, the 16th-century Milanese painter, was clearly aware that his paintings of fruit and vegetables that contained faces when viewed in one orientation look like fruit and vegetables when inverted. More recently, Rex Whistler perfected the art of "reversible faces" (Whistler & Whistler, 1946; see Fig. V.91-3). The curious thing about these faces is that only the upright face can be seen; it is very hard to figure out what the upside-down face is going to look like until the page is turned. These early examples show that upside-down faces are difficult to decode.

Perhaps the modern era of interest in inverted faces stems from seminal work by Yin (1969). This paper proposed that face processing is "special" in that we use "special" brain mechanisms to perceive faces, rather than some general-purpose complex-pattern processor. One key piece of evidence taken to support the "faces are unique" hypothesis is that inverted faces are particularly hard to recognize; it may be true that recognizing inverted buildings and dogs is hard, but nothing suffers as much by inversion as the human face. A very thorough review of this work can be found in Valentine (1988).

Whether faces are really unique remains controversial. One key experiment, by Diamond and Carey (1986), looked at the recognition of human faces and dog faces both by "dog experts" and nonexperts. The results showed that the dog nonexperts (assumed nonetheless to be human experts) suffered greater impairment in the recognition of inverted human faces than inverted dog faces, while the dog experts (again assumed to be human experts too) suffered similar impairment for both sets of stimuli when they were presented upside-down. Diamond and Carey's explanation of their results appealed to expertise. Experts can use holistic cues, while the novice can only look at the individual *features*. Therefore inversion does not prove human faces are unique in the way they are processed—other classes of stimuli may suffer similarly, providing the perceivers are experts in discriminating the class.

This explanation does not stand up to close examination, however; more recent research (e.g., Robbins & McKone, 2007) provides strong support for a "domain-specific" explanation for the effects of inversion. This theory suggests that faces are indeed "special" and the way that we process them is unique. While it is clearly true that inversion affects recognition of many categories of stimuli that have



Figure V.91-1. The original version of the Margaret Thatcher illusion. Viewed upside-down, these two faces look strikingly similar. (From Thompson, P. (1980). Margaret Thatcher: a new illusion. *Perception 9*, 383–4.

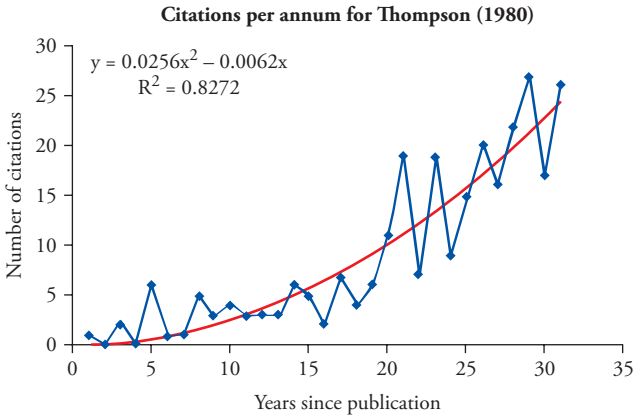


Figure V.91-2. Citations for Thompson (1982). The curve through the data is a second-order polynomial. (Figure created by Peter Thompson.)

a canonical "upright" orientation—faces and text (and dogs!) being obvious examples—nothing compares to the Thatcher illusion in the magnitude of the disruption to our perception caused by this inversion. Perhaps this is why the illusion has had such a lasting appeal.

The difference between holistic encoding of facial information and featural encoding is key to the most widely touted explanation of the Thatcher illusion. Some background may help here. Tanaka and Farah (1993) have shown that we are better at recognizing facial features in situ than in isolation. That is, after studying a face we are asked to decide whether or not a nose belongs to that face. The nose may be presented in isolation or on the face we originally studied. Performance is better when the nose is presented on the face. This suggests that we encode not just the isolated feature of the nose but the whole context



Figure V.91-3. Vegetables by Arcimboldo (left) (The vegetable gardener, Museo Civico Ala Ponzone. Cremona) and "reversible faces" by Rex Whistler (right). (From Whistler & Whistler (1946) ¡OHO! John Lane, the Bodley Head.)

of the nose within the face. It is this "holistic" encoding, it is proposed, that is disrupted by inversion.

If upside-down faces cannot benefit from holistic processes—which operate only on faces in their normal upright orientation—this might explain why inverted faces are harder to recognize. The notion that we rely on both featural and holistic processes to analyze faces has been very widely accepted and with good reason. For example Young, Hellawell, and Hay (1987) demonstrated that a composite face comprising the top half of one familiar face with the bottom half of a second familiar face fuse to form a new face in which it is hard to recognize the two familiar components (see Fig. V.91-4).

Young et al. (1987) explain this by appealing to holistic processes that create the new identity from the stimulus. When the two half-faces are offset one from the other, the

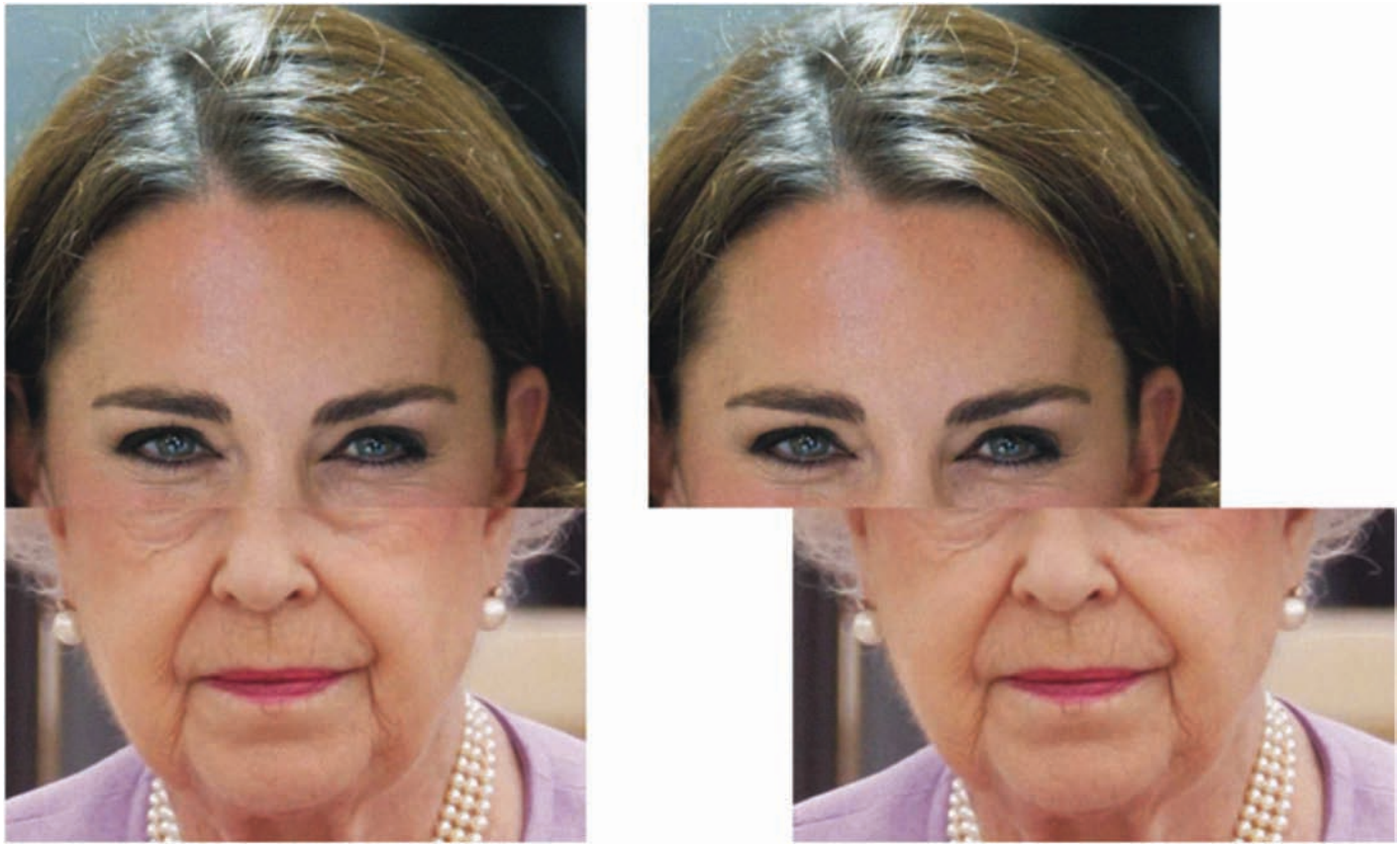


Figure V.91-4. Young et al.'s (1987) demonstration of the effects of holistic encoding. The component identities of the composite face on the left are harder to determine than when holistic processing is disrupted by a lateral shift in the two half images. Catherine, Duchess of Cambridge and Her Majesty Queen Elizabeth II. (Basic Vision [Revised edition] 2012 Snowden, R., Thompson, P. & Troscianko, T. Oxford university Press. Figure 10.12, p. 310.)

recognition of the two components is easier—as holistic processes can no longer fuse the components into a whole new unfamiliar face. Furthermore, viewing the composite faces upside-down actually led to faster recognition of the components than when upright, again, presumably because the inverted face can be analyzed without the interference of holistic processes.

The Thatcher illusion seems to provide some reification of this approach to face processing. When we recognize the identity and expression of a familiar face, we can use both holistic and featural cues. When faces are upside-down, we can no longer use holistic encoding and must resort to the features alone. In the inverted Thatcher face, the features are upright and smiling, and that is how we see the expression.

Unfortunately this explanation leaves us with another problem. If inverted faces are analyzed by their features, then why is it that the "un-Thatcherized" inverted face does not look grotesque—after all, the individual features are all inverted!

One persistent problem with the idea that upside-down faces cannot benefit from holistic encoding is a certain lack of definition. Maurer and colleagues (Mondloch, Le Grand, and Maurer (2002) differentiate between three types of "configural processing": first-order configural processing relies on cues that put two eyes above a nose above a mouth. Second-order configural processing is more subtle and refers to the distances between feature—how far apart the eyes are, how far below the nose the mouth is, and so on. The third type of configural processing—holistic processing—"glues the features into a Gestalt." The Thatcher illusion, as shown in Figure V.91-1, has the eyes, nose, and mouth in exactly the same configuration in the two pictures, which discounts any first-order cues being responsible for the illusion. Furthermore, the eyes are the same distance apart in the two images, the mouth is the same distance from the nose, and so on, discounting any second-order configural involvement. One might argue here that the eyebrows have been moved, but I think this is not critical for two reasons. First, Mrs. Thatcher had very fair eyebrows and so they are not a prominent feature of her face, and second, in some of the other (many) versions of the illusion, the eyebrows are inverted with the eyes whereas in others they are not. A strong argument against Maurer's second-order configural cues being responsible for the effect is a demonstration created by Helmut Leder (Fig. 5.1a in Bruce & Young, 1998), who manipulated the face of Paul Newman; see Figure V.91-5. This only leaves us with Mondloch et al.'s more narrowly defined holistic cues responsible for the illusion, and even this has problems.

Calder, Young, Keane, and Dean (2000) extended the Young et al. (1987) paradigm to study facial expressions. In a composite face comprising, for example, an angry top half with a happy lower half, participants were slower to identify the expressions in the two components when the two halves were aligned (allowing for holistic processing of the whole new face) than when the components were misaligned. Once again inverting the stimuli disrupted this effect, but identification of expression was significantly harder with the inverted faces. So is the Thatcher illusion the result of the fact that facial expressions are difficult to identify in upside-down faces? The evidence is against this theory. First, McKelvie (1995) has shown that upside-down expressions can be recognized, albeit with some difficulty. Second, and more damning, if this theory were true then we should find it difficult to name the expression in Thatcher's inverted expression. And this is not the case. It is plain to see that she is smiling. In addition, she is clearly smiling in both the Thatcherized and the un-Thatcherized versions. It is not the case that we cannot tell the expression of the upside-down, Thatcherized face—it is that we get the expression *wrong*.

VARIANTS AND RELEVANT PARAMETERS OF THE EFFECT

There is nothing particularly special about Margaret Thatcher in the Thatcher illusion. The illusion does and has worked well with any number of other faces. Of course eyes that slant downward and a mouth revealing the upper teeth all help. Nor, it has been claimed, is there anything special in us, the viewers. Apparently rhesus monkeys too are susceptible to the Thatcher illusion (Adachi, Chou, & Hampton, 2009; Dahl, Logothetis, Bulthoff, & Wallraven, 2010). However, these claims must be regarded with caution—it may be that rhesus monkeys fail to discriminate between inverted normal and Thatcherized faces, but this does not say anything about how they perceive the facial expressions. Attempts to demonstrate the effect in pigeons (Jitsumori & Yoshihara, 1997) and baboons (Parron & Fagot, 2008) have both failed.

The criticisms of claims to show the Thatcher illusion in monkeys hold good for prelingual human infants too. Bertin and Bhatt (2004) have shown that six-month-olds can discriminate Thatcherized faces when they were viewed upright but not when they were inverted, but again nothing can be determined about the appearance of the faces to the infants. This adult-like behavior by the infants is taken by Bertin and Bhatt to illustrate that infants are "sensitive to second-order relational information while processing faces." (p. 431). A similar conclusion is reached by Rouse, Donnelly, Hadwin, and Brown (2004), who found autistic children were just as susceptible to the illusion as typically developing children. This study had the advantage of actually asking the children to identify the "unusual" face.

CONCLUSIONS

The appeal of the Thatcher illusion does not seem to be abating more than 30 years after its publication. One clear reason for this is that it is just a wonderful, compelling illusion that makes us laugh and—one hopes—think. Some feel that there is little left to understand about the effect. For example, Rouse et al. (2004, p. 1247) wrote confidently: "The Thatcher illusion is known to result from the computation of second-order relational features, since 'thatcherised' inverted faces are rated as being much less grotesque than upright 'thatcherised' faces (Bartlett & Searcy, 1993)." More recently, Bruce and Young (2012), the doyens of face perception, were somewhat more circumspect:

> Explaining why the upside-down Thatcher images look so alike is tricky (Thompson, Anstis, Rhodes, Jeffery, & Valentine, 2009). Clearly, the relationship between the

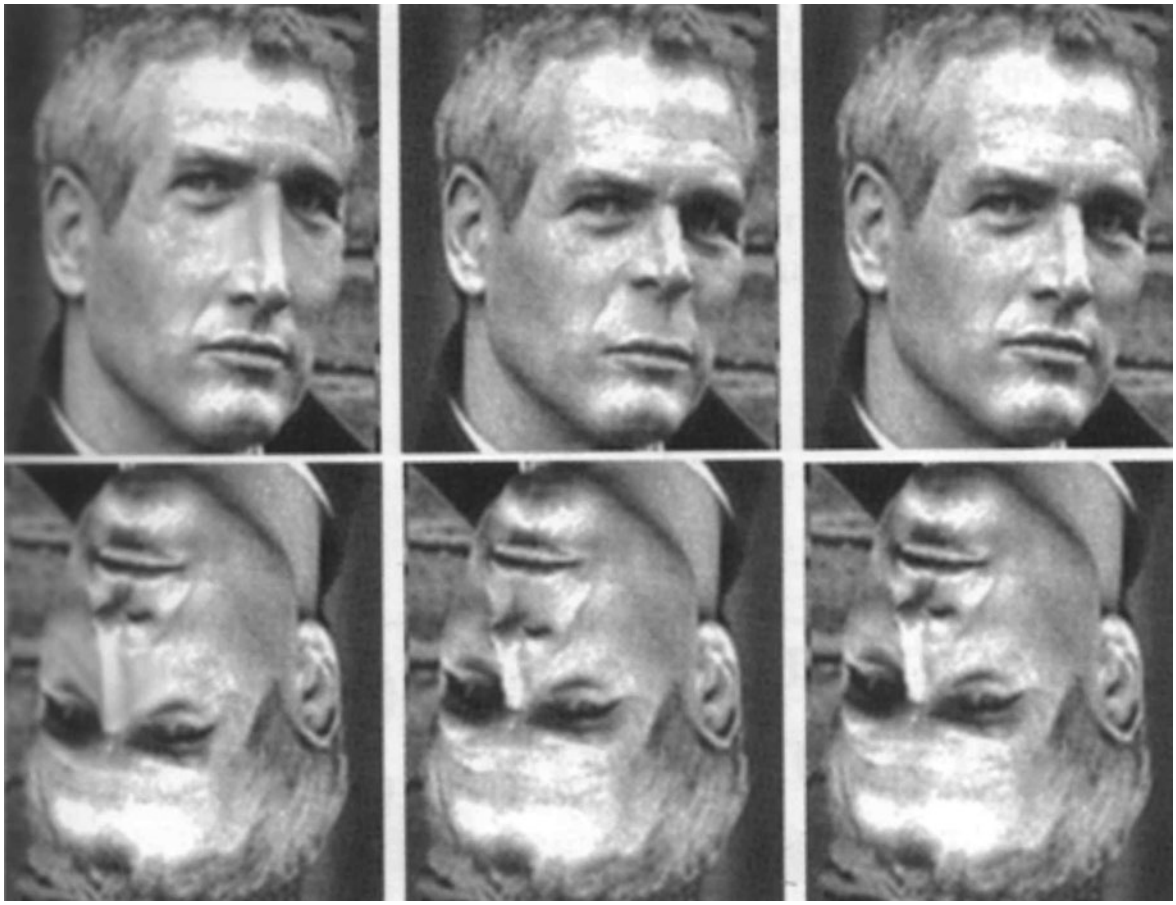


Figure V.91-5. Helmut Leder's demonstration that second-order configural cues, readily apparent in upright faces, are much harder to spot in inverted faces. (Figure created by Helmut Leder). From Bruce and Young (1998). *In the Eye of the Beholder: The science of face perception*. Fig. 5.1(a) p152 Oxford University Press.

> different face features which is so apparent in the upright head is almost invisible in the inverted one. But why? The standard explanation is that we can't see configural information in upside-down faces. . . . It seems to be on the right lines, but it is incomplete. For example, other upside down expressions can be recognised with a bit of difficulty (McKelvie, 1995), but the Thatcherised expression just disappears. The standard counter to this point is that we still have some ability to process upside-down features, and this explains why well-known expressions can be identified in an inverted face. Yet the difference between the features in the upside-down Thatchers (the mouths especially) is substantial. So why can't we see it more easily? We really don't properly understand this. (pp. 185–186)

Perhaps in another 30 years we will understand the illusion.

ACKNOWLEDGMENTS

I would like to thank Arthur Shapiro and Dejan Todorović for useful comments on an earlier draft and particularly Andy Young and Tim Andrews for their many patient attempts to get me to understand something about face processing.

REFERENCES

Adachi, I., Chou, D. P., & Hampton, R. R. (2009). Thatcher effect in monkeys demonstrates conservation of face perception across primates. *Current Biology*, *19*, 1270–1273.

Bertin, E., & Bhatt, R. S. (2004). The Thatcher illusion and face processing in infancy. *Developmental Science*, *7*, 431–436.

Bruce, V., & Young, A. (1998). *In the eye of the beholder: The science of face perception.* Oxford: Oxford University Press.

Bruce, V., & Young, A. (2012). *Face perception.* New York: Psychology Press.

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 527–551.

Dahl, C. D., Logothetis, N. Bulthoff, H. H., & Wallraven, C. (2010). The Thatcher illusion in humans and monkeys. *Proceedings of the Royal Society B*, *277*, 2973–2981.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, *115*, 107–117.

Jitsumori, M., & Yoshihara, M. (1997). Categorical discrimination of human facial expressions by pigeons: A test of the linear feature model. *The Quarterly Journal of Experimental Psychology Section B*, *50*, 253–268.

McKelvie, S. J. (1995). Emotional expression in upside-down faces: Evidence for configurational and componential processing. *British Journal of Social Psychology*, *34*, 325–334.

Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception*, *31*, 553–566.

Parron, C., & Fagot, J. (2008). Baboons (*Papio papio*) spontaneously process the first-order but not second-order configural properties of faces. *American Journal of Primatology*, *70*, 415–422.

Robbins, R., & McKone, E. (2007). No face-like processing for objects-of-expertise in three behavioural tasks. *Cognition*, *103*, 34–79.

Rouse, H., Donnelly, N., Hadwin, J. A., & Brown, T. (2004). Do children with autism perceive second-order relational features? The case of the Thatcher illusion. *Journal of Child Psychology and Psychiatry*, *459*(7), 1246–1257.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *46A*, 225–245.

Thompson, P. (1980). Margaret Thatcher: A new illusion. *Perception, 9*, 383–384.

Thompson, P., Anstis, S., Rhodes, G., Jeffery, L., & Valentine, T. (2009). Thompson's 1980 paper. *Perception*, *38*, 921–932.

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology*, *79*, 471–491.

Whistler, R., & Whistler, L. (1946). *¡Oho!* London: John Lane/The Bodley Head.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*, 141–145.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configural information in face perception. *Perception*, *16*, 747–759.

Chapter 92
The Mona Lisa Gaze Effect

Dejan Todorović

THE PHENOMENON

The Mona Lisa gaze effect is the phenomenon in which certain portraits appear to direct their gaze at observers almost regardless of where observers are located with respect to the picture. An illustration of the effect is presented in Figure V.92-1. This phenomenon is well known but has not been studied much (see Boyarskaya & Hecht, 2012; Goldstein, 1987; Hecht, Boyarskaya, & Kitaoka, 2014; Kitaoka, 2014; Maruyama, Endo, & Sakurai, 1985; Rogers, Lunsford, Strother, & Kubovy, 2003; Todorović, 2006, 2009).

The gaze effect was mentioned as early as the 2nd century AD in a treatise on optics by the famous astronomer and geographer Claudius Ptolemy, who lived in Alexandria, Egypt (Lejeune, 1989; Smith, 1996).[1] A thorough description of the effect was provided by the influential 15th-century German theologian and philosopher Nicolaus Cusanus, who called such images "icons of god" because, godlike, they look at everyone at the same time. He mentioned examples of several contemporary works in which the "subtle art of painting" causes the effect and sent an actual picture of this type to monks in a Benedictine monastery, with specific instructions on how to do display it, where the observers should be located, how they should move, and what they should attend to, including control conditions, somewhat similar to a modern psychophysical experiment.[2]

It is instructive to compare portraits and real three-dimensional (3D) heads in this regard. If a 3D "looker" (a person with fixed head and gaze, or a realistic bust) looks at an observer, the observer can evade the looker's gaze by moving laterally for a short distance; conversely, if the looker does not look at the observer, the observer is generally able to move to a spot from which he or she will meet the looker's gaze. In contrast, with a Mona Lisa–type portrait, the observer will hardly be able to escape the portrait's gaze, almost wherever he or she moves; conversely, if a portrait appears not to look at the observer but, say, to the right, the observer will hardly be able to catch the looker's gaze, because, if he or she moves to the right, the portrait's gaze will appear to move still more to the right. The first type of portraits can be said to have "egocentric gazes," whereas the gazes of other portraits can be called "egophobic."

The key difference between 3D heads and two-dimensional (2D) portraits with respect to the motion of the observer is the manner of transformation of the projected view of the face. With 3D heads, as the observer moves, the projected view changes in characteristic ways, in that some features of the face appear and others disappear from view, such as an ear or a side of the nose; also, projections of some features increase and others diminish in size, such as the views of the two cheeks. On the other hand, the features of a drawn, painted, or photographed face can neither appear nor disappear with the change of the observer's vantage point, although the projection of the face as a whole will change with the angle of projection.

AN UNSATISFACTORY EXPLANATION

The difference between 3D and 2D faces provides the basis for an explanation of the Mona Lisa gaze effect, which I call the "dynamic observer hypothesis": the observer notes that while he or she moves, the projection of the observed face does not transform in the manner that the projection of a real 3D face would transform and thus concludes that the face must have rotated such that it keeps the same projection and thus continue to gaze at him or her.

There are several problems with this explanation. For example, motion is not necessary to document the Mona Lisa effect. In Cusanus's demonstration, observers are stationary and look simultaneously at the same picture from different locations, and they all have the impression that the portrait looks at them. Motion serves only to demonstrate to a single observer how the gaze of the portrait follows him or her as he or she moves about. Moreover, the effect can also be demonstrated to a single stationary observer by presenting simultaneously several copies of the same picture displayed at different angles, which all appear to look at that observer, as illustrated in Figure V.92-1. Furthermore, the dynamic observer hypothesis can at most explain why a portrait would *continue* to gaze at the observer during his or her motion but is silent as to why it would appear to gaze at the observer in the first place, before he or she even starts to move. Finally, whereas the projection of a 3D face rotating strictly in step with the motions of the observer would indeed stay constant, the projection of a 2D face does change when the observer changes location, as

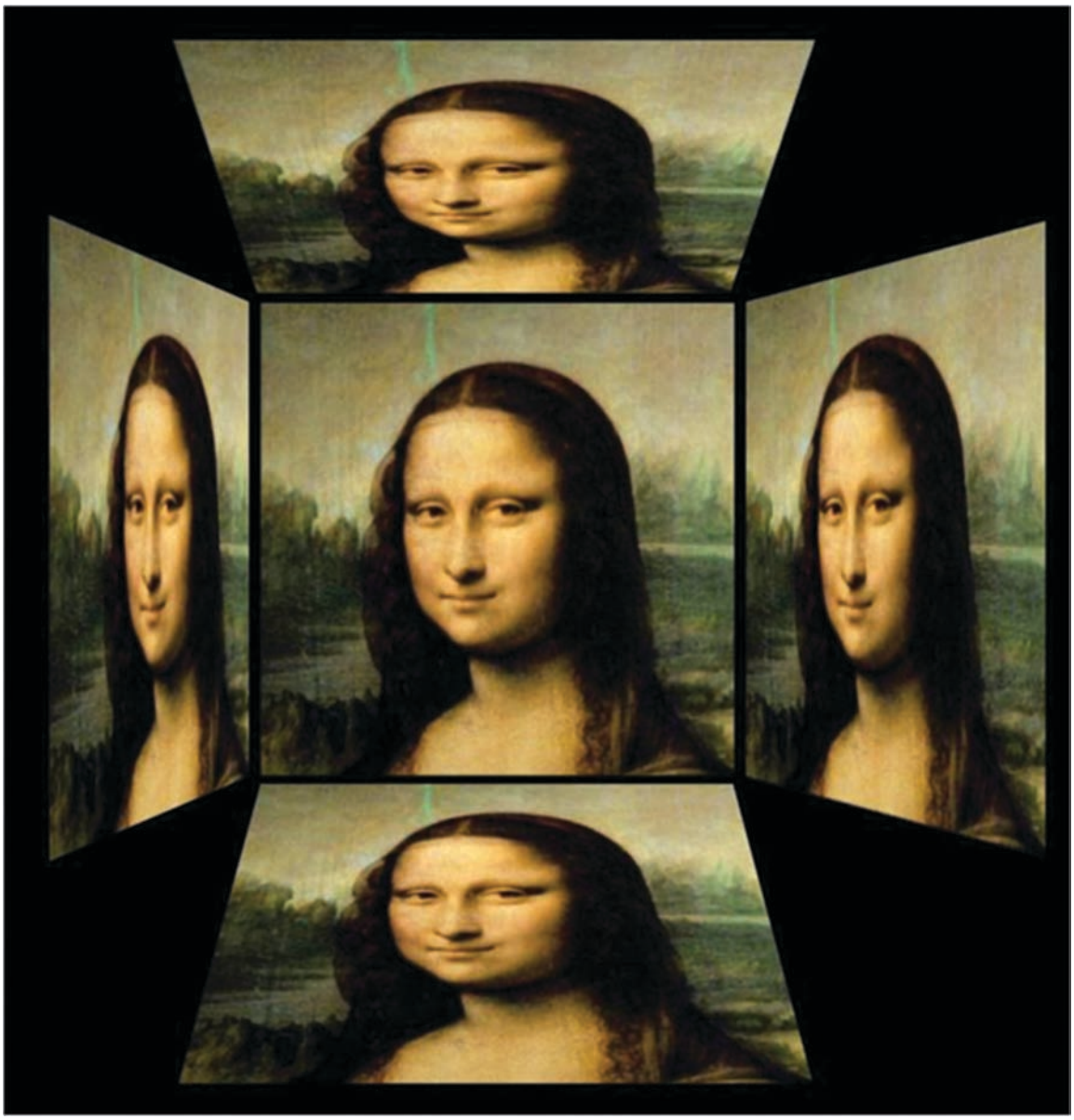


Figure V.92-1. An illustration of the Mona Lisa gaze effect. The image of the Mona Lisa is displayed at various slants as it would project from different angles of observation. These manipulations induce visible deformations of her face but have little effect on the perceived direction of her gaze, which is in all cases directed approximately at the observer. (Author's representation, based on public domain images.)

can be seen in Figure V.92-1, which demonstrates shearing and perspective types of transformations.

PERCEPTION OF GAZE DIRECTION

Perceiving gaze direction is an important social skill, because the direction of gaze of a person or an animal is an indicator of the direction of his or her attention and potential action. So when do we have the impression of being looked at? Still more generally, how do we perceive the gaze direction of a looker, regardless of whether that person's gaze is directed at us or at some other object? A popular notion on how perception of gaze direction works is an intuitively plausible idea that I call the "iris eccentricity hypothesis": it claims that the perceived direction of gaze depends on the position of the iris (and the pupil) within the sclera in the eye opening of the looker in the following way: when the irises are left shifted/centered/right shifted, then the direction of gaze is leftward/straight ahead/rightward, respectively (see Todorović, 2006, for some references essentially endorsing this idea).

However, although the iris eccentricity hypothesis is adequate for faces frontally oriented with respect to the observer, it is not true for slanted faces. To illustrate this, consider the set of portraits in Figure V.92-2, in which some faces confirm the iris eccentricity hypothesis but others disconfirm it. For example, the portraits in the middle row of the matrix all confirm the hypothesis: in the portrait on the left the irises are left shifted and it looks leftward; in



Figure V.92-2. A 3 × 3 matrix of renaissance portraits with factorially combined iris positions and head orientations. Column organization according the position of the irises: left column: irises left shifted; middle column: irises centered; right column: irises right shifted. Row organization according to the orientation of the heads with respect to the observer: top row: head leftward; middle row: head frontal; bottom row: head rightward. For easier comparison, the portraits are manipulated such that that their interocular distances are approximately equal and their heads are upright. (Author's representation, based on public domain images.)

the portrait in the middle the irises are centered and it looks straight ahead at the observer; and in the portrait on the right the irises are right shifted and it looks rightward. Furthermore, in the top left portrait the irises are left shifted and it looks leftward, and in the bottom right portrait the irises are right shifted and it looks rightward. However, the remaining portraits disconfirm the hypothesis. In the middle portrait in the top row the irises are centered but it looks leftward, and in the middle portrait in the bottom row the irises are also centered but it looks rightward. Furthermore, in the leftmost portrait in the bottom row the irises are left shifted but it gazes at the observer, and in the rightmost portrait in the top row the irises are right shifted but it also gazes at the observer. These examples indicate that the perceived direction of gaze does not depend only on the positions of the irises but also on the orientation of the head of the portrait with respect to the observer, as described next.

Several studies have confirmed that the perception of gaze direction is affected by the head turn of the looker (Anstis, Mayhew, & Morley, 1969; Cline, 1967; Conty, Tijus, Hugueville, Coelho, & George, 2006; Doi & Ueda, 2007; Gibson & Pick, 1963; Kluttz, Mayes, West, & Kerby, 2009; Langton, 2000; Langton, Watt, & Bruce, 2004; Otsuka, Mareschal, Calder, & Clifford, 2014; Ricciardelli & Driver, 2008; Todorović, 2006; West, 2013). An early

direct demonstration of the fact that perception of gaze direction does not depend only on the eyes was provided by Wollaston (1824).[3] In the illustration of Wollaston's effect in Figure V.92-3a, the original portrait of Mona Lisa, whose irises are right shifted, is on the right, and on the left her head is mirrored except for the eyes, which are simply switched but remain right shifted. We can see that although the two portraits have the same positions of irises, the one on the right gazes approximately at the observer, whereas the one on the left gazes rightward.

Perception of gaze direction, including the Mona Lisa effect and the Wollaston effect, can be studied using not only realistic portraits but also very simplified, schematic faces (Maruyama et al., 1985; Todorović, 2009). In the illustration in Figure V.92-3b, the schematic face on the left and the schematic face on the right have identical eyes, and yet the one on the right appears to gaze approximately at the observer, whereas the on the left appears to gaze rightward, similar to the two portraits in Figure V.92-3a. In the composite display in the middle of Figure V.92-3b, the two schematic faces are partially overlapped, sharing the same pair of eyes. If, while observing this display, we direct our attention to its right half, and thus mentally "extract" the right head from the composite and ignore everything outside of

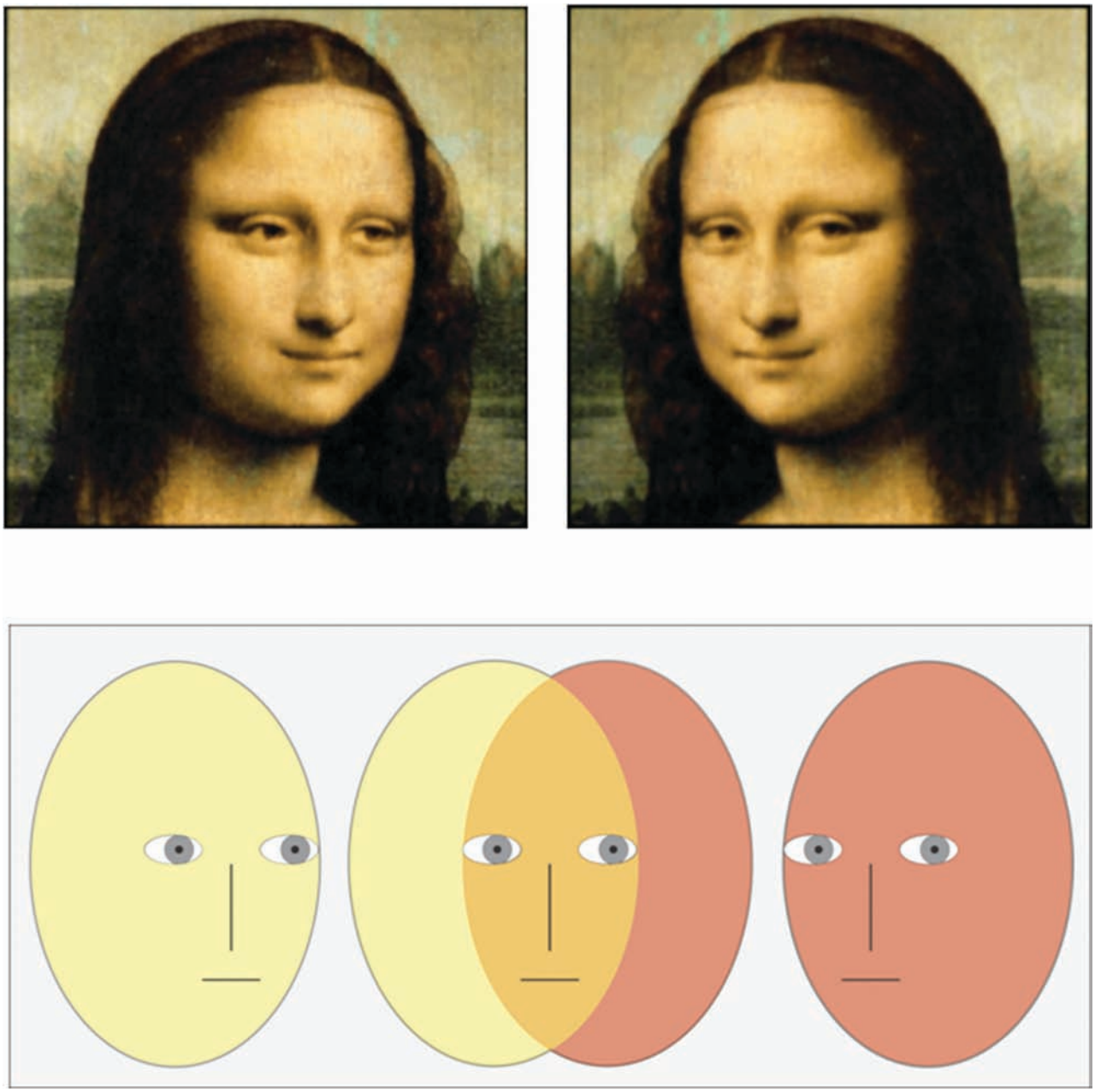


Figure V.92-3. Illustrations of the Wollaston effect. Faces with identical eyes gaze in different directions. (a) Realistic faces. (b) Schematic faces. (Author's representations, based on public domain images.)

it, it will appear to gaze approximately at us, exactly as the single schematic face on the right of Figure V.92-3b; on the other hand if, in a similar fashion, we pay attention to the left head in the composite display, it will appear to gaze rightward, just as the schematic face on the left.

The difference between the two Mona Lisa portraits in Figure V.92-3a is that in the original portrait the head appears to be turned leftward with respect to the observer, whereas in the mirror version the head appears to be turned rightward. On the other hand, the two schematic faces in Figure V.92-3b differ in what I call "face eccentricity" (Todorović, 2009), which is the relative position of the inner facial features (eyes, nose, mouth) within the head outline. It is plausible that, in schematic faces, face eccentricity functions as a crude but strong cue of head orientation. Thus for both realistic and schematic faces the differences in perceived gaze directions, with iris position held constant, seems to be due to the differences in perceived head orientations. This fact is sometimes described as an illusion-inducing influence of head orientation, shifting perceived gaze direction away from the "true" gaze direction signaled by the positions of the irises. However, a simple analysis of the geometry of the situation, described next, indicates that this is a mischaracterization and that the perception of the direction of gaze in the Wollaston effect is not an illusion but precisely what is to be expected from a mechanism that attempts to correctly detect gaze direction with respect to the observer.

Note that what the position of the irises ("iris eccentricity") indicates is not gaze direction per se but gaze direction *relative to the looker*, that is, the orientation of the eyes with respect to the head of the looker; for example, right-shifted irises (from the observer's perspective) indicate that the gaze of the looker is directed to the *looker's* left. On the other hand, gaze direction *relative to the environment* (including the observer) depends not only on iris eccentricity but also on the orientation of the looker's head. The portraits in Figure V.92-2 provide illustrations of this fact. In the middle portrait in the middle row the head is frontally oriented and the irises are centered, and it appears to gaze straight at the observer. In the portraits to the left/right in the same row, the heads are also frontal but the eyes are turned left/right, and their gaze directions are also turned left/right. However, the same shifts in gaze direction can also be produced by keeping the irises centered and turning only the heads left/right, as illustrated by the top/bottom portraits in the middle column. This is a trivial consequence of the fact that the eyes are in the head; therefore, even when they are fixed with respect to the head, their orientation in space will change when the orientation of the head changes. Thus these portraits represent situations in which the gaze directions of the lookers are straight ahead relative to their own heads but averted with respect to observers and are correctly seen as averted. When the eyes are turned in one direction and the head is turned in the opposite direction, then gaze direction can be directed at the observer, as illustrated in the bottom left and top right portrait.

Gaze direction with respect to the environment is an imaginary vector directed from the fovea of the looker to the gazed-at object, passing through the center of the pupil perpendicularly to the cornea. This direction might conceivably be detected by the observer on the basis of some indicators of the curvature of the cornea at the pupil's center. However, the reliability of such indictors is questionable, even when observing the eyes of 3D lookers from a relatively small distance, and they are hardly available when judging gaze directions of portraits. A plausible alternative source of information is based on the fact that gaze direction G can be represented as $G = E + H$, that is, as the sum of eye turn E (gaze direction with respect to the looker) and head turn H (turn of head with respect to the observer; see Todorović, 2006, 2009). Available information about eye turn and head turn of lookers is much more robust than information about the curvature of the transparent cornea. Carlin, Calder, Kriegeskorte, Nili, and Rowe (2011) have used this equation to define gaze directions of computer-generated faces and have found that subjects can successfully retrieve gaze directions from such images. Although these authors argue for a head-invariant representation of gaze direction in the anterior superior temporal sulcus, it is clear from the previously described decomposition that in the detection of gaze direction head turn does play a role; that is, the observer cannot rely solely on the information about the position of the irises of the looker but must also take into account the orientation of the looker's head with respect to the observer.

Whereas actual eye turns and head turns are specified in angular terms, the corresponding *perceived* attributes must be retrieved from some visual indicators of these angles. Based on results of experiments with schematic heads, Todorović (2009) has suggested that perceived gaze direction (PG) can be modeled as a linear combination of iris eccentricity (IE) and face eccentricity (FE):

$$PG = iw * IE + fw * FE$$

Here PG is perceived gaze direction with respect to the observer, IE and FE are visual indicators of eye turn and head turn, and iw and fw are empirically determinable weights of these indicators.

Note that the equation $PG = 0$ corresponds to the situation in which the gaze of the looker appears to be directed at the observer. When this is true, the previous algebraic equation has two types of solutions. The first type corresponds to the condition in which both eccentricities are zero; that is, when $IE = FE = 0$. The meaning of this is that the gaze of the looker will be perceived to be directed at the observer if the irises are centered in the sclera ($IE = 0$) and the head has frontal orientation with respect to the observer ($FE = 0$). This situation, which is denoted as the "zero eccentricities condition," is exemplified by the central portrait in Figure V.92-2, which indeed appears to saliently look at the observer.

The second type of solution corresponds to the condition in which neither eccentricity is zero. This solution can be expressed as follows:

$$IE / FE = -(fw / iw) = C$$

This is not a single solution but a family of solutions. This condition states that the ratio of the numerical measures of eye turn and head turn is equal to a certain negative (given that the weights are positive) constant value C, which is

specified by the ratio of the weights. This means that the gaze will be perceived to be directed at the observer whenever the turn of the eyes in one direction is accompanied by a proportional turn of the head in the other direction (hence the negative sign). Such a situation, which is called the "counterbalanced eccentricities condition," corresponds to the portrait of the Mona Lisa and the bottom left and top right portraits in Figure V.92-2. When the head is turned only slightly to one side, the eyes will need to be turned only slightly to the other side to maintain direct gaze; the more the head is turned, the more the eyes will have to be turned to counterbalance the head turn and keep the gaze directed at the observer. The reader can check this by looking in the mirror and turning the head from side to side as if gesturing "no" while keeping the gaze directed at oneself (Todorović, 2006). This simple exercise is also a good demonstration that perception of gaze direction cannot depend only on the position of the iris/pupil within the sclera: note that during head motion the relative position of the iris within the eye opening shifts continuously from one end to the other, whereas the gaze stays directed at the observer all the time, because the orientation of the eyes with respect to the environment is kept constant due to the coordination of head and eye motions.

Finally, when only one eccentricity is zero and the other is not zero, the equation does not have a solution (*PG* cannot be zero), meaning that the gaze cannot be directed at the observer. Such cases are represented in the middle portraits in the top and bottom row and left and right column in Figure V.92-2.

EXPLAINING THE MONA LISA EFFECT

Assuming that perception of gaze direction is based on signals of eye turn and head turn, the Mona Lisa effect can be accounted for as follows. Consider first portraits exhibiting the zero eccentricities condition. The top part of Figure V.92-4a depicts two observers looking at such a portrait. Observer A stands directly opposite the picture, so that his line of sight is more or less perpendicular to its plane, whereas observer B stands off to the side of the picture, with an oblique line of sight. Observer A registers that the face of the portrait is symmetrical, which is plausibly a strong cue that the looker's head is frontally oriented with respect to him, as sketched in the bottom part of the figure. Furthermore, he registers that the irises of the portrait are centered. These two pieces of information specify that the gaze of the portrait is directed at him. Observer B registers the same two pieces of information. First, the face is symmetrical. This corresponds to the situation that the looker's head is frontally oriented with respect to her, as sketched in the bottom part of the figure; I discuss this aspect in more detail later. Second, she also registers that the irises of the portrait are centered. Therefore, the two pieces of information specify that the gaze of the portrait is directed at her. It follows that both A and B will have the impression that the portrait is gazing at them.

Figure V.92-4b depicts a portrait exhibiting the counterbalanced eccentricities condition, with observers in same positions as in Figure V.92-4a. In this case the face of the portrait is not symmetrical, indicating that the orientation of the head with respect to the observer is not frontoparallel. The problem of head pose estimation was studied extensively in the computer vision literature (see the survey by Murphy-Chutorian & Trivedi, 2009), whereas the question how human observers detect head orientation has not been explored much (see Lawson, Clifford, & Calder, 2011; Poppe, Rienks, & Heylen, 2007; Todorović, 2006; Troje & Siebeck, 1998; Wilson, Wilkinson, Lin, & Castillo, 2000). There are several potential 2D cues for head orientation such as face eccentricity, the angle of the nose, relative visibility of the two cheeks, degree of departure from symmetry, and the like. It can be assumed that based on such cues observer A registers the orientation of the head of the portrait as deviating from frontoparallel and not being perpendicular to his own line of sight. He also registers that the irises deviate from the center of the eye to a certain extent. This situation is represented in the bottom part of the picture. Given that, by assumption, the two deviations are counterbalanced, observer A will have the impression that the portrait gazes at him. Observer B is presented with much the same cues for head turn and eye turn as observer A. Thus for her the orientation of the portrait's head in virtual pictorial 3D space will deviate from *her* frontoparallel orientation to a similar extent as for observer A, as depicted in the bottom portion of the figure. Furthermore, she will register much the same deviation of the irises from the center of the eyes. Therefore, similarly to observer A, the two deviations will be in counterbalance for observer B as well, meaning that she also will have the impression that the portrait gazes at her. In sum, according to the presented considerations, portraits exhibiting either the zero eccentricities condition or the counterbalanced eccentricities condition will show the Mona Lisa effect; that is, they will convey the impression of gazing at observers for many different observer positions with respect to the image.

This account is simplified to some extent, and certain caveats and qualifications are in order. For one, humans have two eyes. and when they look at a nearby object their gaze directions are not parallel but convergent; this aspect was not taken into account in the previous analysis. Another aspect that was ignored was that the sclera is a curved surface in 3D, so that the 2D projections of the indicators of eye turn (location of the iris/pupil within the sclera) depend to some extent on the orientation of the eyes of the looker with respect to the observer (see Langton, 2011). In addition, there is a difference between the two observers in the examples in Figure V.92-4a, in that for observer A the orientation of the physical plane of the portrait is frontoparallel, which agrees with the symmetry of the face, whereas for observer B this plane is oblique, which is incongruous with facial symmetry, making the situation somewhat ambiguous. Furthermore, because of the lateral view, for observer B the projection of the image as a whole will be somewhat compressed compared with the projection for observer A, as illustrated in Figure V.92-1. Similar considerations apply for Figure V.92-4b. Lateral views also involve differential foreshortening; that is, the portions of an image further from the observer will be compressed more than the nearer portions, and this effect

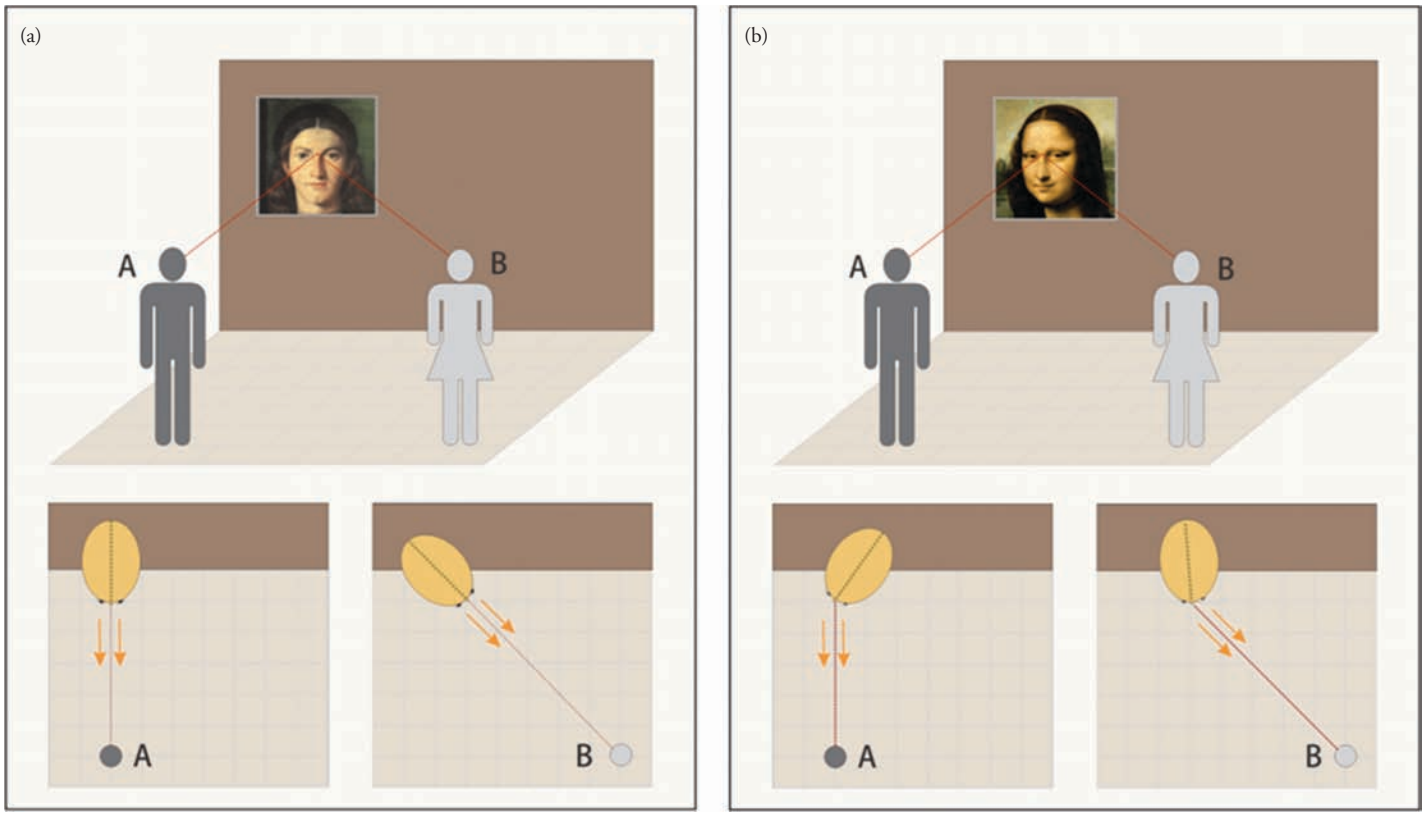


Figure V.92-4. Explanation of the Mona Lisa effect. (a) Fully symmetrical portrait. (b) Counterbalanced asymmetrical portrait. Top parts: Side view. Bottom parts: Top view of observer and hypothesized position of the looker's head in virtual pictorial space. Red lines: Lines of sight. Green dashed lines: Axes of heads. Orange arrows: Gaze directions. (Author's representations, based on public domain images.)

will be stronger for increasing angles of view. One consequence of this is that the half of the face of the portrait further away from the observer will be foreshortened to a somewhat larger degree than the nearer half, therefore changing somewhat the 2D cues for the orientation of the head and thus the predicted gaze direction. I calculated the degree of this distortion for the conditions in an experiment in which schematic faces were presented to subjects on screens with two orientations, one frontoparallel and the other turned by 30° (Todorović, 2009); although it was of modest size, the predicted effect was roughly in agreement with the obtained difference of perceived gaze direction in the two situations, amounting to a small failure of the Mona Lisa effect, on the order of 2° to 3°. However, for more extreme oblique views such effects may be substantial and lead to a complete breakdown of the Mona Lisa effect. Such a breakdown for more extreme angles of observation was indeed demonstrated by Hecht et al. (2014).

NOTES

1. Here is what Ptolemy wrote about the effect:

> It is also assumed that the image of a face painted on panels follows the gaze of [moving] viewers to some extent even though there is no motion in the image itself, and the reason is that the true direction of the painted face's gaze is perceived by means only of the stationary disposition of the visual cone that strikes the painted face. The visual faculty does not recognize this, but the gaze remains fixed solely along the visual axis, because the parts themselves of the face are seen by means of corresponding visual rays. Thus, as the observer moves away, he supposes that the image's gaze follows his. (Smith, 1996, p. 124)

This explanation is not particularly clear. For an interpretation by the editor of Ptolemy's work, see Smith, (1996, p.124, footnote). One possible reason for the difficulty is that the English text (for a French translation of this passage, see Lejeune, 1989, p. 79) is a translation from Latin, which was translated from a lost Arabic text, which may or may not have been directly translated from Ptolemy's lost original, which was written in Greek.

The fact that Ptolemy mentions that faces were painted on panels is intriguing. Around that time in Egypt there was a flourishing production of so-called Fayum mummy portraits, images skilfully and naturalistically painted on wooden panels that were attached to covered faces of mummified bodies, probably representing the deceased. It may be speculated that in such circumstances the effect of vivid, piercing eyes appearing to follow the observer's every step was particularly noticeable.

2. Cusanus wrote:

> Fix this [icon of god] anywhere, say at the northern wall. You, brothers, position yourselves around the icon, not far away, and look at her! Everyone among you, from wherever

> he looks at her, will have the experience that he alone is gazed at by her. To the brother who stands in the east it will seem that this face looks at east, and to the brother who stands in the south it will seem that the face looks at the south, and to the one who stands in west that the face looks at west. At first you will wonder how it is possible that the icon looks at everybody and at every single one of you at the same time. Because the imagination of the one who stands in the east cannot comprehend it at all that the gaze of the icon is directed at some other area, such as towards west or south. Then the brother who stands in the east may move towards west; and he will have the experience that the gaze is directed at him, as it was before in the east. And since he knows that the icon is fixed and has not been changed, he will wonder about the change of the unchangeable gaze. And if he goes from west to east, fixing his look all the time at the icon, he will experience that the gaze of the icon all the time goes along with him. (Simon, 2004, pp. 59–60).

Cusanus does not seem to have offered an explanation as to how the "subtle art of painting" might cause the wondrous effect. Rather, his interest in the phenomenon was theological:

> So, observing brother, approach the icon of god! Place yourself first in the east, then in the south, and finally in the west! And since the gaze of the icon looks at you in the same way from all sides and does not leave you wherever you turn, you will be prompted into contemplation. You will be stimulated to say: Lord, now I see in this picture Your providence in a sort of a sensory experience . . . I plan to raise you, beloved brothers, by way of such a sensory appearance, through a special devotional exercise, up to the mystic theology. (von Kues, 1967, my translation, Simon, 2004, p. 63)

3. Wollaston's paper is available on the web and is well worth reading. However, the modern PDF format still cannot recover an ingenious detail of the original journal publication in 1824. Wollaston's figures involve two portraits with identical top parts, containing the eyes, and different bottom parts, containing the nose, mouth, and so on, depicted in different orientations. The manner in which the identity of the top parts is demonstrated to the reader is as follows. One full-size journal page contains a complete portrait, composed of the top part and the bottom part. Another page immediately precedes this page in the journal volume, but its top part is cut away at a slant, and it contains only the bottom part of the face, different from the bottom part on the immediately succeeding full-size page. When this cut-away page is placed over the full-size page, one complete portrait is seen, and when it is turned, another complete portrait is seen, with a different bottom portion. Thus by turning the cut-away page back and forth, readers can unambiguously and convincingly convince themselves that the eyes are indeed identical in the two portraits but that they appear to gaze in two different directions.

REFERENCES

Anstis, S. M., Mayhew, J. W., & Morley, T. (1969). The perception where a face or television "portrait" is looking. *The American Journal of Psychology*, *82*, 474–489.

Boyarskaya, E., & Hecht, H. (2012). Differentiating the differential rotation effect. *Acta Psychologica*, *140*, 252–265.

Carlin, J. D., Calder, A. J., Kriegeskorte, N., Nili, H., & Rowe, J. B. (2011). A head view-invariant representation of gaze direction in anterior superior temporal sulcus. *Current Biology*, *21*(21), 1817–1821.

Cline, M. G. (1967). The perception of where a person is looking. *The American Journal of Psychology*, *80*, 41–50.

Conty, L., Tijus, C., Hugueville, L., Coelho, E., & George, N. (2006). Searching for asymmetries in the detection of gaze contact versus averted gaze under different head views: A behavioural study. *Spatial Vision*, *19*, 529–545.

Doi, H., & Ueda, K. (2007). Searching for a perceived stare in the crowd. *Perception*, *36*, 773–780.

Gibson, J. J., & Pick, A. D. (1963). Perception of another person's looking behavior. *The American Journal of Psychology*, *76*, 386–394.

Goldstein, E. B. (1987). Spatial layout, orientation relative to the observer, and perceived projection in pictures viewed at an angle. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 256–266.

Hecht, H., Boyarskaya, E., & Kitaoka, A. (2014). The Mona Lisa effect: Testing the limits of perceptual robustness vis-à-vis slanted images. *Psihologija*, *47*(3), 287–301.

Kitaoka, A. (2014). Left bias of gaze perception in a cartoon face. *Psihologija*, *47*(3), 315–318.

Kluttz, N. L., Mayes, B. R., West, R. W., & Kerby, D. S. (2009). The effect of head turn on the perception of gaze. *Vision Research*, *49*(15), 1979–1993.

Langton, S. R. H. (2000). The mutual influence of gaze and head orientation in the analysis of social attention direction. *The Quarterly Journal of Experimental Psychology*, *53A*, 825–845.

Langton, S. R. H. (2011). Gaze perception and visually mediated attention. In R. B. Adams Jr., N. Ambady, K. Nakayama, & S. Shimojo (Eds.), *The science of social vision* (pp. 108–132). Oxford: Oxford University Press.

Langton, S. R. H., Watt, R. J., & Bruce, V. (2004). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences*, *4*, 50–59.

Lawson, R. P., Clifford, C. W., & Calder, A. J. (2011). A real head turner: Horizontal and vertical head directions are multichannel coded. *Journal of Vision*, *11*(9):17, 1–17.

Lejeune, A. (1989). *L'optique de Claude Ptolémée*. Leiden: Brill.

Maruyama, K., Endo, M., & Sakurai, K. (1985). An experimental consideration on "Mona Lisa gaze effect." *Tohoku Psychologica Folia*, *44*, 109–121.

Murphy-Chutorian, E., & Trivedi, M. M. (2009). Head pose estimation in computer vision: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *31*(4), 607–626.

Otsuka, Y., Mareschal, I., Calder, A. J., & Clifford, C. W. (2014). Dual-route model of the effect of head orientation on perceived gaze direction. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(4), 1425–1439.

Poppe, R., Rienks, R., & Heylen, D. (2007). Accuracy of head orientation perception in triadic situations: Experiment in a virtual environment. *Perception*, *36*, 971–979.

Ricciardelli, P., & Driver, J. (2008). Effects of head orientation on gaze perception: How positive congruency effects can be reversed. *The Quarterly Journal of Experimental Psychology*, *61*(3), 491–504.

Rogers, S., Lunsford, M., Strother, L., & Kubovy, M. (2003). The Mona Lisa effect: Perception of gaze direction in real and pictured faces. In S. Rogers & J. Effken (Eds.), *Studies in perception and action VII* (pp. 19–24). Hillsdale, NJ: Lawrence Erlbaum.

Simon, H. (2004). Bildtheoretische grundlagen des neuzeitlichen bildes bei Nikolaus von Kues. *Concilium Medii Aevi*, *7*, 45–76.

Todorović, D. (2006). Geometrical basis of perception of gaze direction. *Vision Research*, *46*, 3549–3562.

Todorović, D. (2009). The effect of face eccentricity on the perception of gaze direction. *Perception*, *38*(1), 109–132.

Troje, N. F., & Siebeck, U. (1998). Illumination-induced apparent shift in orientation of human heads. *Perception*, *27*, 671–680.

von Kues, N. (1967). De visione Dei. In L. Gabriel (Ed.), *Philosophisch-theologische schriften. Bd. 3* (pp. 93–219). Wien, Germany: Herder.

Watts, P. M. (1982). *Nicolaus Cusanus: A fifteenth-century vision of man*. Studies in the History of Christian Thought 30. Leiden: Brill.

West, R. W. (2013). The effect of head turn and illumination on the perceived direction of gaze. *Perception*, *42*, 495–507.

Wilson, H. R., Wilkinson, F., Lin, L., & Castillo, M. (2000). Perception of head orientation. *Vision Research*, *40*, 459–472.

Wollaston, W. H. (1824). On the apparent direction of eyes in a portrait. *Philosophical Transactions of the Royal Society of London*, *114*, 247–256.

Chapter 93
The Wobbling Face Illusion

Sayako Ueda and Akiyoshi Kitaoka

SEEING DOUBLE

Here we report a novel type of illusion, the wobbling face illusion, which was first proposed by Suga (2003) and quantitatively examined and named by Ueda, Kitaoka, and Suga (2011; see also Martinez-Conde & Macknik, 2010). This illusion is easily reproduced by duplicating some of the features (eyes or mouth) of a portrait and placing the copied features close to the originals.

Figure V.93-1a shows an example of this illusion. Some observers can perceive the face image as being a four-browed, four-eyed, two-nosed, and two-mouthed woman. Others cannot. Since the image causes the brain to try to integrate the duplicated features, viewers may experience an unstable, wobbling feeling that is akin to double vision. It should be noted that double vision or diplopia is a troublesome visual disorder with a variety of medical causes, while the wobbling face illusion is not a disorder.

We examined how many people do and do not see this illusion by asking students at Ritsumeikan University (Kyoto, Japan) and Japan Women's University (Kawasaki, Japan) to rate their experience of the illusion. The results showed that 82% of the students reported the illusion (with 44% reporting a weak illusion), while 18% of the students did not. This finding suggests that the majority of observers see this illusion, while a minority of about 20% do not. The cause of these individual differences remains an open question.

The motif of the wobbling face may possibly be traced back 3,000 years to the mask of a god used in a royal ceremony of the ancient Chinese Zhou (周) dynasty, which arose in the 11th century BC and was overthrown in 256 BC. The god Fang Xiang Shi (方相氏, Hososhi in Japanese) played the role of protecting the palace and graves of the royal family from demons or evil spirits and is always represented as a mask with four gold eyes (Fig. V.93-1b). This mask looks like a stimulus that demonstrates the wobbling face illusion.

The idea of Hososhi was accepted by the Japanese royal family as early as 706 AD, according to the ancient Japanese formal document *Shoku-Nihongi* (Ueda, 1988). The four-eyed mask of Hososhi can be seen annually in the traditional event *Tsuinashiki* in some Shinto shrines in Japan. According to the priests at Yoshida shrine, one such shrine in Kyoto, two of the four eyes are used to see things normally while the other two are supposed to see invisible or evil things, and the four eyes watch over the entire world. The relationship of the purpose of Hososhi's duplicated eyes and the effect of the wobbling face illusion is not clear, but we imagine that Japanese people's responses to this illusion may be partly based on their familiarity with Hososhi.

In the 20th century, American avant-garde photographer Man Ray created *Marquise Casati* (1922), a double-eyed portrait of a woman that resembles the wobbling face illusion. In a textbook of visual illusions, Komachiya (2005) described this work as being "illusional." The wobbling face illusion was explicitly noted by Suga (2003), who referred to the effect as "eyes gravitation." The popular website Freaking News has maintained a gallery of many photographic portraits of famous persons that have been processed to produce this illusion and interest (http://www.freakingnews.com/Double-Vision-Pictures--1762-0.asp). In 2010 these pictures were introduced in *Scientific American* (Martinez-Conde & Macknik, 2010) as illustrating "illusions involving eyes." The "double face" illusion is another name for the wobbling face illusion, used by Hancock and Foster (2012) in their investigative report. The results of their study indicated that there were wide individual differences in the ability to detect the wobbling face when presented briefly and subsequently masked. In a recognition task, duplication of facial parts disturbed the identification of celebrities' faces, though the disturbance was much weaker than that shown in face inversion. Moreover, they reported that participants were quicker to judge whether or not the face was doubled as it was rotated away from upright. Furthermore, they showed that normal and wobbling faces did not pop out from each other but revealed a processing cost of 40 to 60 ms per wobbling face.

Interestingly, this phenomenon occurs not only in photographs but also in illustrations (Fig. V.93-1c). Readers may notice a similarity to a technique used in cartoons to express motion (Fig. V.93-1d). However, there are some differences between these images and the wobbling face illusion. In cartoons representing motion, the whole face is duplicated. This duplication represents the movement afterimage. On the other hand, in wobbling face images, duplication is not perfect; it is limited to facial features,



Figure V.93-1. Examples of (a) the wobbling face illusion, (b) Hososhi, (c) the illustrated wobbling face, (d) the cartoon art form. (a) and (b) were reprinted from Ueda, Kitaoka and Suga, Wobbling appearance of a face induced by doubled parts, Perception-London, 2011, with permission from SAGE Publications.

and the contour is not duplicated. In this way, this partial duplication does not resemble the movement afterimage.

Why are wobbling face images unsteady and difficult to perceive? What prevents some people from perceiving faces with four eyebrows, four eyes, two noses, and two mouths?

EVIDENCE FOR FACE-SPECIFIC PROPERTIES

Is the wobbling face illusion specific to the face? The human face conveys a great variety of information, such as an individual's identity, age, gender, race, and emotional state. We are typically able to recognize upright faces more quickly and more accurately than other types of visual stimuli (Yin, 1969) and can remember large numbers of individual faces over long periods (Bahrick, Bahrick, & Wittilinger, 1975). When we look at pictures, our gaze tends to focus on faces more often than on non-faces (Yarbus, 1967). If the wobbling face illusion is specific to the face, some facial processing or neural mechanisms should underlie this illusion.

Martinez-Conde and Macknik (2010) suggested that the unstable feeling may arise because of the special nature of facial processing. If viewers can perceive a face with four eyebrows, four eyes, two noses, and two mouths, there should be no unstable feeling. This suggests that the duplicated facial features might prevent neurons for face recognition from processing this visual information correctly.

We recently presented evidence for face-specific properties of the wobbling face illusion (Ueda et al., 2011). This study involved two experiments. In experiment 1, we tried to determine the nature of the unstable feeling for the wobbling face. We created original stimuli by using colored photographs of a Japanese female model as a facial object and a house as a nonfacial object. Stimuli were created in three conditions: (a) no duplication, (b) partial duplication (for face: eyes; for house: windows), and (c) partial duplication in mirror-reverse (Fig. V.93-2a). Computer graphic techniques were used to control variables. These stimuli were shown in upright or inverted orientations. This enabled us to ascertain if facial processing is involved as a cause of the unstable feeling, because face inversion is known to impair face perception and recognition by disturbing the integration of face configuration (Bartlett & Searcy, 1993; Freire, Lee, & Symons, 2000; Leder & Bruce, 2000; Rhodes, Brake, & Atkinson, 1993; Searcy & Bartlett, 1996).

In experiment 1, significantly more observers reported the illusory feeling for face images (59%) than for house images (2%) and for upright images (33%) than for inverted images (28%). Moreover, for face images, most observers reported the illusory feeling for duplication images (78%) and for duplication in mirror reverse images (100%), and no one reported the illusory feeling for no-duplication images. We therefore concluded that the wobbling face illusion is specific to the face and dependent upon facial processing.

The most frequent reports of observers who experienced the illusory feeling for duplicated face images included "I couldn't perceive a four-eyed woman" and "I was not sure which I should focus on, the upper eyes or the lower ones." On the other hand, only two participants who did not experience the illusory feeling reported that they perceived a four-eyed woman. Participants who experienced the illusory feeling for the facial image in mirror-reverse most often reported "I couldn't detect her gaze direction" or "I detected either averted right eyes or averted left eyes." These reports suggest that the difficulty of stable perception in the wobbling face illusion might arise from some competition or rivalry between the upper pair of eyes and the lower pair of eyes. To investigate this idea, we performed an additional experiment using the same participants.

In experiment 2, the duplicated parts of each stimulus were changed in size or shape in order to change the impressions of pairs of parts. For the face image, the duplicated lower eyes were made larger. Bigger and lower-placed eyes are characteristic of babyish faces (Enlow, 1982). For the house image, the duplicated lower windows were given Japanese characteristics (Fig. V.93-2b). These stimuli were prepared in upright and inverted orientations.

We found that most observers easily formed a single impression of the upright house image (95%), the inverted house image (100%), and the inverted facial image (80%), but only a few observers formed a single impression of the upright facial image (10%). In the upright facial image, observers who reported this illusion estimated age for the upper eyes, when they were shown without the lower eyes, as significantly

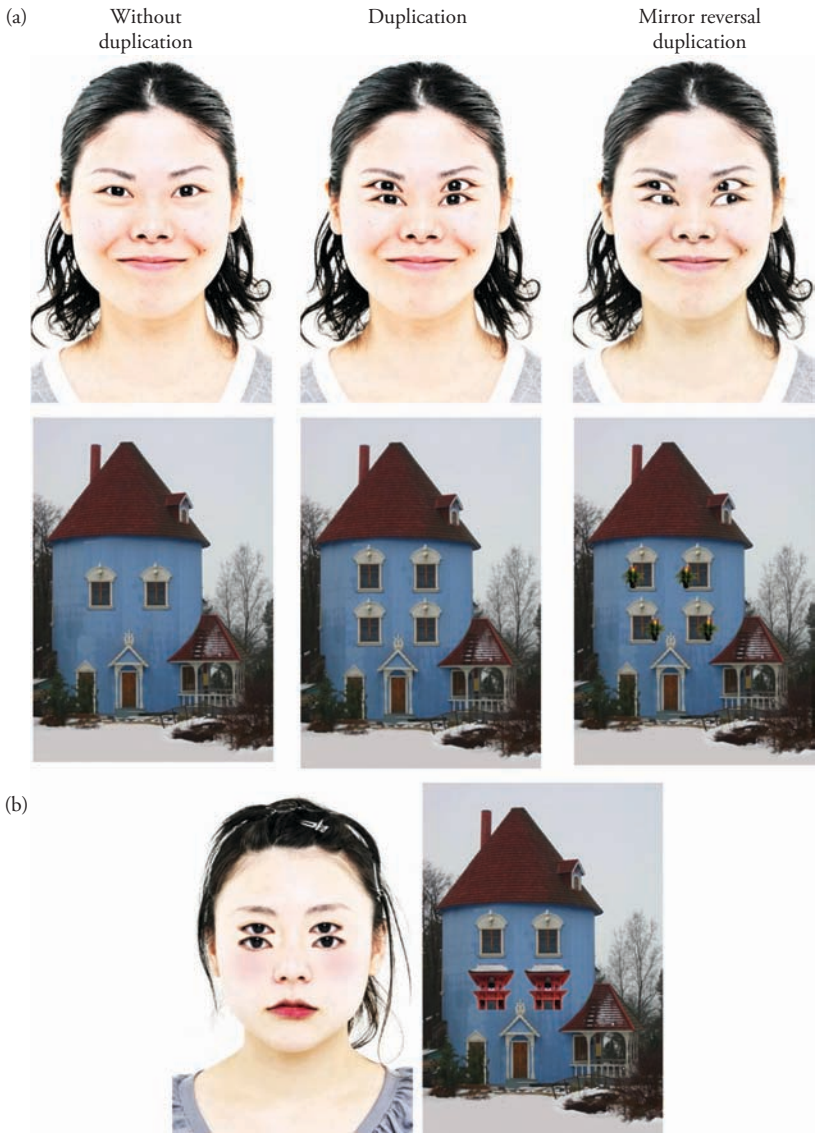


Figure V.93-2. (a) Stimuli for experiment 1. (b) Stimuli for experiment 2. Reprinted from Ueda, Kitaoka and Suga, Wobbling appearance of a face induced by doubled parts, Perception-London, 2011, with permission from SAGE Publications.

older (mean 24.56 years old) than the lower ones, when they were shown without the upper eyes (mean 15.72 years old). Meanwhile, those who reported no illusory experience and observed a four-eyed face rated its age at 19.00 years old, which was the average of the illusion group's two age ratings. Therefore, we concluded that the wobbling face illusion may be attributable to facial processing competition between the upper pair of eyes and the lower pair of eyes.

In sum, experiments 1 and 2 indicated that the wobbling face illusion should be specific to the face, and the unstable feeling may arise from rivalry of perception between doubled facial features.

WHICH PART CREATES THE STRONGEST ILLUSION?

The wobbling face illusion may differ in magnitude for different facial parts. To investigate this question, we conducted experiment 3 using the following five stimuli (Fig. V.93-3a): (a) without duplication, (b) eyebrow duplication, (c) eye duplication, (d) nose duplication, and (e) mouth duplication. These stimuli were shown in upright or inverted orientations.

The upright face produced a significantly stronger illusion than the inverted one, and the unstable feeling was evoked by the following parts, in order of strength: eyes > mouth > nose = eyebrows. That is, eyes had the strongest effect on this illusion. We imagine that this eye dominance might have caused the mask of Hososhi to be equipped with double eyes in ancient days. Moreover, this property can probably explain why eyes and mouth are preferred as doubled parts, as demonstrated on the Freaking News website (http://www.freakingnews.com/Double-Vision-Pictures--1762-0.asp).

The simplest explanation for this eye dominance may be that the eyes are attended to more than other facial features on a variety of tasks (e.g., Althoff & Cohen, 1999; Henderson, Williams, & Falk, 2005; Luria & Strauss,



Figure V.93-3. (a) Stimuli for experiment 3. (b) Interindividual wobbling face. (c) Interage wobbling face. (d) Interexpressional wobbling face. (e) Intersexual wobbling face. In (b), (c), and (e), all features except the eyes and mouths were averaged shapes. (These images were originally created by Sayako Ueda.)

1978; Walker-Smith, Gale, & Findlay, 1977; Yarbus, 1967). Moreover, the eyes are central to social communication because they indicate emotions, direction of attention, and intentions (Baron-Cohen, 1995; Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997).

Based on these findings, we tried to demonstrate additional variations of the wobbling face illusion in which the features are not exact duplicates of each other, shown in Figures V.93-3b–e. Figure V.93-3b, which demonstrates an interindividual wobbling face, was created by combining Audrey Hepburn's and Marilyn Monroe's faces. This face suggests that the wobbling face illusion can be produced by mixing different faces. Figure V.93-3c, which demonstrates an interage wobbling face, was created by combining the girlhood and adulthood faces of Elizabeth Taylor and shows that the wobbling feeling may arise from seeing a face of varying ages. Figure V.93-3d displays an interexpressional wobbling face and was created by combining a neutral face and a smiling one. It shows that the wobbling face illusion may be independent of the emotion that is expressed and may be created by neutral, smiling, and even mixed expressions. Figure V.93-3e, which shows an intersex wobbling face and was created by combining a female face and a male face, shows that we may see female, male, or sometimes androgynous faces. Moreover, in Figure V.93-3b, observers may also see a "wobbling gaze" because the two eyes have different gaze directions.

IS THE WOBBLING FACE ILLUSION A MOTION ILLUSION?

A variety of motion illusions are observed in stationary images (for reviews, see Kitaoka & Ashida, 2007; Pinna & Spillmann, 2005). For some motion illusions, eye movements have been examined in terms of the motion of retinal images or retinal slips (e.g., Pinna illusion: Pinna & Brelstaff, 2000; visual-delay illusion: Kitaoka & Ashida,

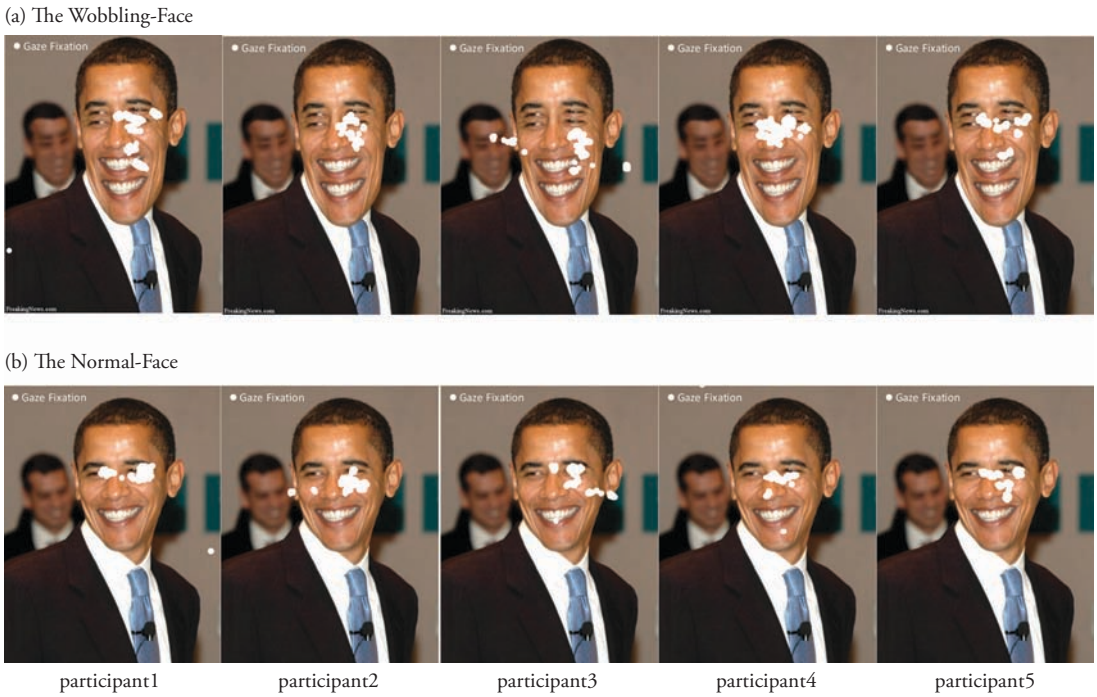


Figure V.93-4. Eye-gaze patterns: (a) Wobbling face image. (b) Original normal face image. Data of gaze fixation is shown as white circles. Facial parts duplication didn't derive special eye movement. Reprinted from Ueda, Kitaoka and Suga, Wobbling appearance of a face induced by doubled parts, Perception-London, 2011, with permission from SAGE Publications.

2007) or eye movement instabilities (rotating snakes illusion: Kuriki, Ashida, Murakami, & Kitaoka, 2008).

If eye movements cause rivalry of perception between doubled facial parts, the wobbling face illusion can simply be attributed to a motion illusion in stationary images. In order to examine the possibility of eye movement instabilities, we examined eye movements in five observers who reported illusory feelings while observing the wobbling face image and the original face image of US President Barack Obama, which was presented by Martinez-Conde and Macknik (2010, p. 38).

We found no special eye movements that shifted back and forth between the duplicated parts. This suggests that the unstable perception did not reflect retinal slips or eye movement instabilities. In addition, when the head was slowly moved, the unstable feeling of the wobbling face did not appear to change. These findings cast doubt on the possibility that the wobbling face illusion is a motion illusion; instead, we speculate that this illusion may depend on a cognitive process, such as attention.

The rivalry of perception between doubled facial parts might arise from the rivalry of attention between the doubled facial parts. The observer's attention may be directed to only one pair of eyes rather than to four eyes at one time, because facial processing might impose a limit on the number of facial parts that can be processed at once. How can we explain this limit? It is possible that internal facial templates are involved. Their existence was suggested by observation of the newborn face preference phenomenon (e.g., Fantz, 1963). The internal facial templates may be constructed from facial parts found in normal faces: two eyebrows, two eyes, one nose, and one mouth. If facial processing is assumed to be biased to attend to these internal facial templates in order to detect normal faces, responses should be strong to a duplicated facial part, and competition should arise between doubled facial parts. Such a template may also explain why this illusion is specific to the face.

However, duplicating facial features does not always induce the wobbling impression. Figures V.93-5a and b show examples. Figure V.93-5a, which was created by duplicating the irises of the eyes, demonstrates that the wobbling feeling is not induced; instead, the feeling arises that the irises are fluctuating from right to left or up and down. Further, there seems to be a difference in strength of the feeling between vertical duplication and horizontal duplication. Figure V.93-5b, which was created by duplicating the nostrils, produces the feeling that the nostrils are flaring. These images may demonstrate phenomena different from the wobbling face illusion, but the partial duplication of these features may not be strong enough to invoke the neural mechanisms for the wobbling face illusion. Therefore, we distinguish three types of duplicating effect: (a) the wobbling effect, which is characterized by the unnatural perceiving of objects; (b) the fluctuating effect, which is characterized by the natural perceiving of objects; and (c) the static effect, which is characterized by the static perceiving of objects.

We conjecture that two factors affect the wobbling face illusion. One factor is configural processing, which integrates the local processing of each feature. Face processing has often been suggested to depend largely on configural processing (Diamond & Carey, 1986; Yin, 1969; Young, Hellawell, & Hay, 1987). Duplication of features may have

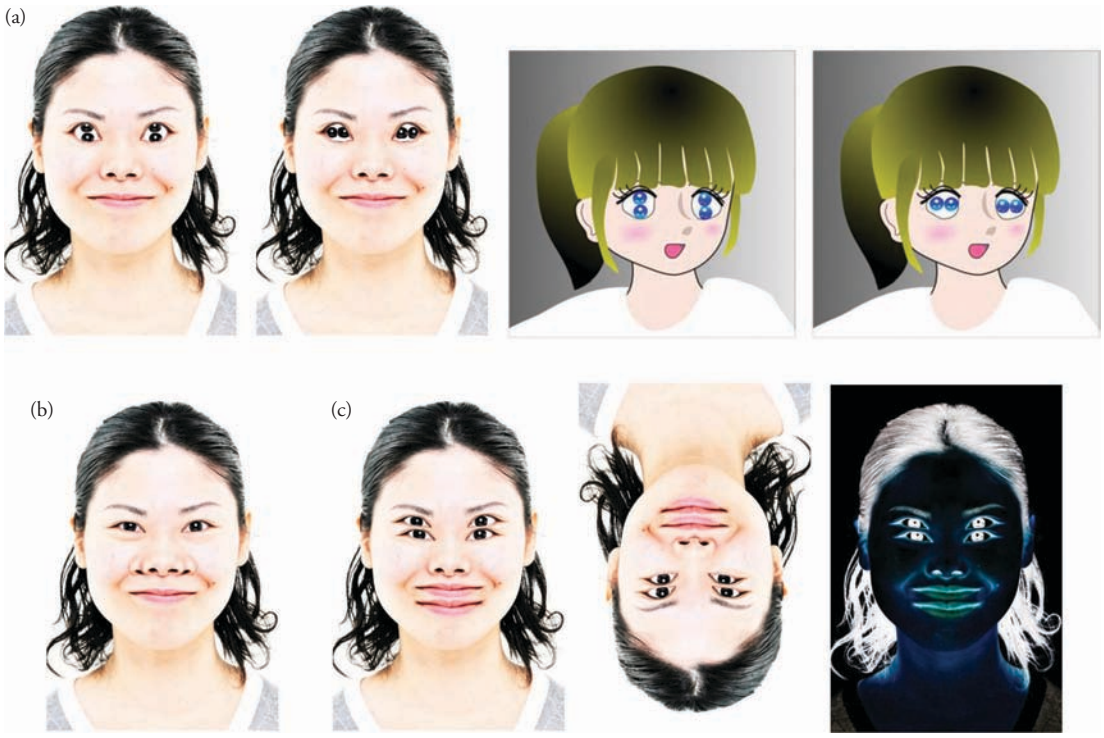


Figure V.93-5. (a) Iridic duplication. (b) Nostril duplication. (c) The wobbling face, its inversion, and its photographic negative version.

some influence on this typical configural processing in some cases but not in others. Configural processing is thought to be particularly disrupted by inversion and photographic negatives (for reviews, see Maurer, Le Grand, & Mondloch, 2002). If the onset of the wobbling effect is closely linked to configural processing, the wobbling feeling may decrease in inversion and photographic negatives. In fact, as Figure V.93-5c shows, these disruptive transformations may decrease the wobbling feeling.

The other factor that may affect the wobbling face illusion is the plasticity or movability of facial parts. Face-based biological motion is assumed to be either rigid, in which the head translates and rotates but does not change shape, or nonrigid, in which the expression changes but the head does not move. These types of motion may be processed independently and may make different contributions to the perception of social information. Irises and nostrils are essentially readily movable parts, and their motions are regarded as nonrigid motion. Nonrigid motions of internal facial features provide visual information relating to speech, expression, and eye movement (O'Toole, Roark, & Abdi, 2002). Because of the importance of these forms of social information, it may be easier for us to detect them. We suggest that the wobbling face illusion occurs when duplication of facial parts creates a rigid motion trajectory.

CONCLUSIONS

A novel phenomenon, the wobbling face illusion has been examined and discussed in terms of face perception and motion illusion, as well as attention. It is suggested that this illusion is face-specific, it is not a simple motion illusion that depends on eye movements, and it may be based on attention instability. Moreover, it resembles a reversible or ambiguous figure such as the Necker cube or the silhouette illusion (Kayahara, 2003; Troje & McAdam, 2010), because they demonstrate that one perspective on a shape suddenly switches to another in the mind of the observer. However, in the wobbling face illusion, the switch between doubled parts occurs more frequently, at much more than 1 cycle/s. This illusion then can be regarded as a new, frequently reversing type of ambiguous figure.

REFERENCES

Althoff, R. R., & Cohen, N. J. (1999). Eye-movement-based memory effect: A reprocessing effect in face perception. *Journal of Experimental Psychology: Learning Memory and Cognition*, *25*, 997–1010.

Bahrick, H. P., Bahrick, P. O., & Wittilinger, R. P. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General*, *104*, 54–75.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: MIT Press.

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or asperger syndrome. *Journal of Child Psychology and Psychiatry*, *38*, 813–822.

Bartlett, J. C., & Searcy, J. (1993). Inversion and configuration of faces. *Cognitive Psychology*, *25*, 281–316.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, *115*, 107–117.

Enlow, D. H. (1982). *Handbook of facial growth* (2d ed.). Philadelphia: Saunders.

Fantz, R. L. (1963). Pattern vision in newborn infants. *Science*, *40*, 296–297.

Freire, A., Lee, K., & Symons, L.A. (2000). The face-inversion effect as a deficit in the encoding of configural information: Direct evidence. *Perception*, *29*, 159–170.

Hancock, P. J. B., & Foster, C. (2012). The "double face" illusion. *Perception, 41*(1), 57–70.

Henderson, J. M., Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory & Cognition*, *33*, 98–106.

Kayahara, N. (2003). Silhouette Illusion. http://www.procreo.jp/labo/labo13.html

Kitaoka, A., & Ashida, H. (2007). A variant of the anomalous motion illusion based upon contrast and visual latency. *Perception, 36*, 1019–1035.

Komachiya, A. (2005). Visual illusion and art. In T. Goto & H. Tanaka (Eds.), *Handbook of the science of visual illusions* (pp. 427–435). Tokyo: University of Tokyo Press. (in Japanese)

Kuriki, I., Ashida, H., Murakami, I., & Kitaoka, A. (2008). Functional brain imaging of the rotating snakes illusion by fMRI. *Journal of Vision, 8*(10):16, 1–10.

Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural information in face recognition. *The Quarterly Journal of Experimental Psychology*, *53A*, 513–536.

Luria, S. M., & Strauss, M. S. (1978). Comparison of eye movements over faces in photographic positives and negatives. *Perception*, *7*, 349–358.

Martinez-Conde, S., & Macknik, S. L. (2010). The eyes have it. *Scientific American Mind*, *20*(1), 36–41.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*, 255–260.

O'Toole, A. J., Roark, D. A., & Abdi, H. (2002). Recognizing moving faces: A psychological and neural synthesis. *Trends in Cognitive Sciences, 6*, 261–266.

Pinna, B., & Brelstaff, G. J. (2000). A new visual illusion of relative motion. *Vision Research, 40*, 2091–2096.

Pinna, B., & Spillmann, L. (2005). New illusions of sliding motion in depth. *Perception, 34*, 1441–1458.

Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition*, *47*, 25–57.

Searcy, J. H., & Bartlett, J. C. (1996). Inversion and processing of component and spatial-relational information of faces. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 904–915.

Suga, T. (2003). *Three types of temperament*. Kyoto, Japan: Kitaoji-Shobo. (in Japanese)

Troje, N. F., & McAdam, M. (2010). The viewing-from-above bias and the silhouette illusion. *i-Perception, 1*, 143–148.

Ueda, S. (1988). Hososhi no shoso. *Bulletin of the Kashihara Archaeological Institute*, *10*, 345–377. (in Japanese)

Ueda, S., Kitaoka, A., & Suga, T. (2011). Wobbling appearance of a face induced by doubled parts. *Perception*, *40*, 751–756.

Walker-Smith, G. J., Gale, A. G., & Findlay, J. M. (1977). Eye movement strategies involved in face perception. *Perception*, *6*, 313–326.

Yarbus, A. (1967). *Eye movements and vision*. New York: Plenum.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *8*, 141–145.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, *16*, 747–759.

Chapter 94
Adaptation Aftereffects in the Perception of Faces

Michael A. Webster

Most people are very adept at recognizing a face they have seen previously, or inferring from the face an individual's traits, such as his or her gender, age, or expression. These abilities suggest that some aspects of the visual representation of faces remain very stable. Yet, surprisingly, face perception may also involve highly dynamic processes that are continuously recalibrated by the variety of faces to which we are exposed. In particular, the appearance of a face can be rapidly and dramatically changed after viewing—and thus adapting—to a different face (Webster & MacLeod, 2011).

Webster and Maclin (1999) explored these adaptation effects by using images of faces that were distorted to produce different facial configurations (Fig. V.94-1a). After viewing a face in which the features were pulled together (middle left), the original face appeared pulled apart (i.e., more like the middle right image). An example of a distortion aftereffect in a highly familiar face is shown in Figure V.94-2 (from Carbon & Leder, 2006), while an animation illustrating the aftereffect is provided on Video V.94-1. The perceived distortions induced by the adaptation are pronounced. For example, in one task observers were asked to judge the range of distortions that looked like a possible photograph of a real person (MacLin & Webster, 2001; Webster & MacLin, 1999). After adapting, the original face no longer appeared plausible, and the range of realistic images was instead again strongly shifted toward the adapting face (to null the aftereffect of a perceptual shift away from the adapting face).

Leopold, O'Toole, Vetter, and Blanz (2001) subsequently showed that similar aftereffects occur for the "distortions" that distinguish actual faces. Their stimuli were based on a morphable face model derived from three-dimensional shape and reflectance measurements from a set of subjects and allowed an individual's face to be defined and varied relative to the average for the set (Fig. V.94-1b). Adaptation to an "antiface" (with features complementary or opposite to an original target face) caused the average face to appear more like the target. In addition to identity, adaptation also biases many of the natural attributes we judge from faces, including invariant properties such as gender or ethnicity (Webster, Kaping, Mizokami, & Duhamel, 2004) or attractiveness (Rhodes, Jeffery, Watson, Clifford, & Nakayama, 2003), as well as variant states such as expression (Hsu & Young 2004; Webster et al., 2004), age (O'Neil & Webster 2011; Schweinberger et al., 2010; O'Neil, Mac, Rhodes, & Webster, 2014), or direction of gaze (Jenkins, Beaver, & Calder, 2006). Thus after adapting to a young female face,

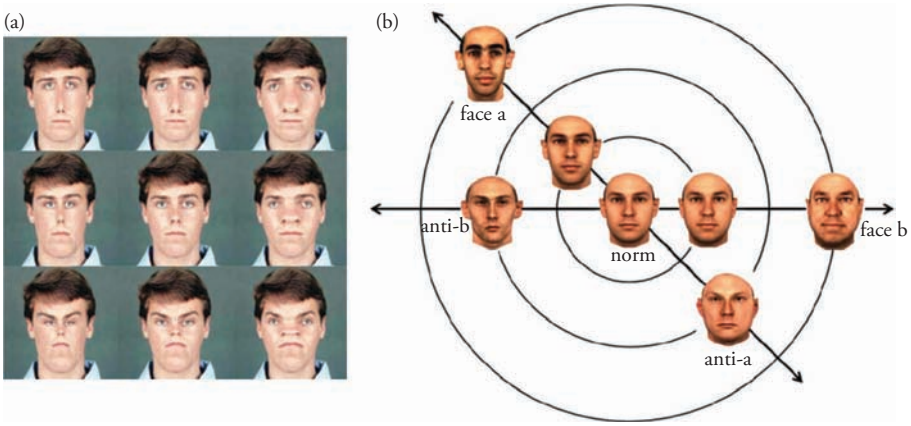


Figure V.94-1. Examples of stimuli used to measure face aftereffects (Webster & MacLeod, 2011). (a) Distortions in an original face image (center) by contracting or expanding the facial configuration (Webster & MacLin, 1999). After adapting to a contracted face, the original face looks too expanded. (b) variations in facial identity by varying the identity strength of faces relative to an average or prototype face (center; Leopold et al., 2001). Adaptation to an antiface (e.g., anti-b) causes the average face to appear more like the opposite, original face (e.g., face b).

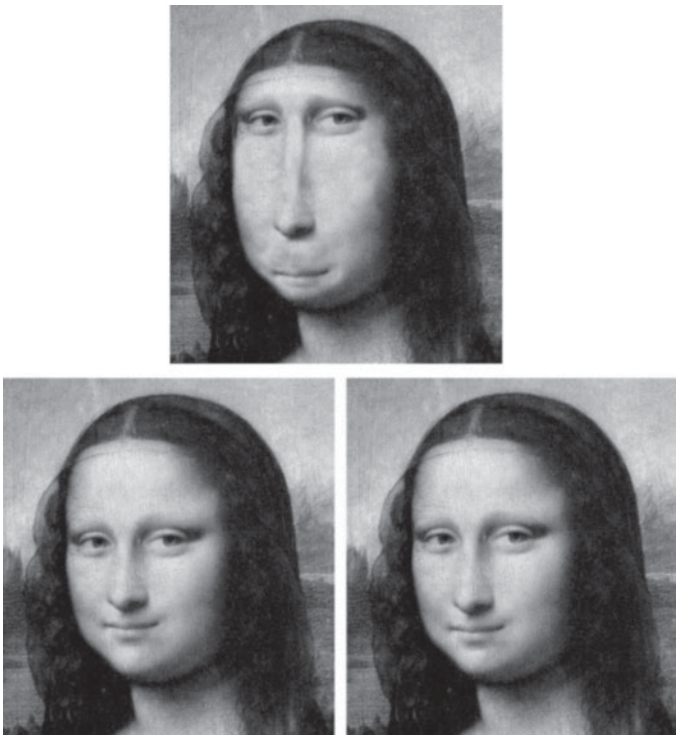


Figure V.94-2. (a) A distorted version of the Mona Lisa (Carbon & Leder, 2006). View the image for a minute or more. Note that the distortion may appear less exaggerated as you continue to view it, consistent with renormalization of the perceived facial configuration. (b) The original portrait image. After adapting to the distortion, the original may appear too contracted and thus unlike the remembered face. (c) A slightly distorted version of the portrait. After adaptation, this may appear more like the original, since the physical expansion nulls the perceptual contraction induced by the adaptation.

a middle-aged face appears older and an androgynous face appears more masculine.

A large number of investigations of face aftereffects exist (Webster & MacLeod, 2011; Strobach & Carbon, 2013).



Video V.94-1. A demonstration of face aftereffects. Exposure to the distorted face causes the original face to appear distorted in the opposite direction.

This work has been important in extending the study of adaptation, which has traditionally been focused on simpler stimulus attributes, to more abstract and potentially higher levels of visual analysis (Webster, 2011, 2015). In this regard, the plasticity of face perception is perhaps more notable for how similar in form it is to the many classic visual aftereffects of color, form, and motion. There are close correspondences between the pattern and dynamics of adaptation across many stimulus dimensions (Clifford, 2002; Leopold, Rhodes, Müller, & Jeffery, 2005), suggesting that common principles of adaptation contribute to the visual encoding of most if not all perceptual attributes (Webster, 2011). Interestingly, if there is something special about adaptation to faces, it may be that there is not a clear spatial analog of the temporal effects. The appearance of most stimuli can be biased by both the prior context (adaptation) and the concurrent surround (simultaneous contrast; Schwartz, Hsu, & Dayan, 2007). For example, a gray square appears lighter on a dark surround, and vertical bars appear tilted clockwise when surrounded by counterclockwise bars. Simultaneous contrast effects have not been reported for faces; that is, an undistorted face does not appear to change when surrounded by distorted faces. This may in part reflect a foveal specialization of face processing, so that faces in the periphery are too weak to generate a contrast effect. Even when a face falls slightly outside the point of fixation, its appearance becomes ambiguous and highly labile. A striking demonstration of this is a recently reported illusion when viewing a rapid sequence of different faces in the periphery (Tangen, Murphy, & Thompson, 2011). These appear grotesquely distorted by the contrast with the preceding face. Moreover, even a single static face can take on multiple and very different appearances, potentially because of the temporal contrasts generated from successive fixations (Simas, 2000).

The perceptual illusions induced by adaptation have traditionally been an important tool for probing visual mechanisms (Graham, 1989), and one reason for the intense interest in face adaptation is that the specific pattern of aftereffects has potential implications for understanding the visual mechanisms encoding faces. Two prominent accounts of face coding are exemplar-based models, in which a face is represented by its absolute values along a set of dimensions; and norm-based models, in which the values define how the face differs from a prototype (Valentine, 1991; Valentine. Lewis, & Hills, 2015). These models differ in whether the average face has a special status in "face space," as well as in the predictions they make for adaptation (Rhodes, Robbins, Watson, Clifford, & Nakayama, 2005; Webster & MacLeod, 2011). In a norm-based code, adaptation to a distinctive face should renormalize the prototype by adjusting responses so that the adapting face itself becomes more neutral. Thus, importantly, adaptation should change the appearance of the face we are adapting to (making it look more average), similar to the way that a color becomes less saturated as we continue to view it. This shift can perhaps be experienced as we continue to inspect the distorted image of Mona Lisa (Fig. V.94-2a), which may appear less distorted with time. Moreover, this model predicts that there should be no aftereffect when adapting to the prototype—since this is simply reinforcing the norm (in the same way that hues do not appear more saturated

after adapting to gray). Conversely, in the exemplar-based model, no face is special. Adaptation to any face, including the average, will reduce sensitivity to that specific exemplar, causing other faces to appear repulsed away from the adapting level. These distinctions have also been considered in terms of whether the dimension of variation is encoded by a set of many narrowly tuned channels (exemplar-based) or a pair of broadly tuned channels coding opposing deviations from the mean (norm-based). A number of studies have pointed to a special role of the norm in face aftereffects, consistent with a norm-based code (Rhodes et al., 2005; Webster & MacLeod, 2011). For example, distorted adapting faces alter the appearance of an undistorted face but not vice versa (Webster & MacLin, 1999), and adapting to an expanded face causes all faces to appear less expanded, rather than a repulsion relative to the specific level of the expansion (Robbins, McKone, & Edwards, 2007). Moreover, adapting to different ages appears to normalize perceived age – after adaptation to a young or old face, all faces including the adapting face may appear older or younger, respectively (O'Neil et al., 2014). However, recent studies have questioned whether adaptation to some facial dimensions such as gender strength shows a pattern consistent with the predictions of a simple renormalization (Storrs & Arnold, 2012; Zhao, Series, Hancock, & Bednar, 2011), and this issue continues to be debated (McKone, Jeffery, Boeing, Clifford, & Rhodes, 2014; Ross, Deroche, & Palmeri, 2014; O'Neil, Mac, Rhodes, & Webster, 2015). It is also possible that both renormalization and repulsion can occur but involve distinct neural adjustments (Müller, Schillinger, Do, & Leopold, 2009).

Of course whether adaptation can reveal the structure of face space depends in part on whether the sensitivity changes arise at neural levels where the stimulus is actually encoded as a face. Thus a further major interest has been in trying to isolate and probe aftereffects that might arise at face-specific stages. This is not straightforward, for adaptation alters responses throughout the visual pathway, and thus higher stages can inherit the pattern of response changes that occur at earlier levels (Kohn, 2007; Webster, 2015). For example, expression aftereffects can be induced by simple curves (Xu, Dayan, Lipkin, & Qian, 2008) or orientation gradients (Dickinson, Almeida, Bell, & Badcock, 2010) that do not themselves look like faces. Yet a number of results suggest that at least part of the aftereffect arises from adaptation at high if not face-specific levels of processing. The aftereffects show strong transfer across changes in size or retinal position (Afraz & Cavanagh, 2009; Leopold et al., 2001; Zhao & Chubb, 2001). Moreover, distortion aftereffects can transfer across rotations in the face and thus are tied to the object rather than the spatial orientation (Watson & Clifford, 2003). Identity aftereffects also transfer between face images that are uniformly stretched—a distortion that changes the low-level features and their spacing but preserves the perceived identity (Hole, 2011). High-level aftereffects have also been implicated by findings that the aftereffects depend on the perceived category rather than the physical similarity of the face images (Bestelmeyer et al., 2008) and that they are strongly reduced in the absence of awareness (Moradi, Koch, & Shimojo, 2005). However, perhaps the most obvious test of a face-specific aftereffect has been surprisingly difficult to demonstrate. Faces are much harder to recognize when they are upside down, and this inversion effect is considered a hallmark of face processing. Yet face aftereffects can be nearly as strong for inverted as upright faces (Susilo, McKone, & Edwards, 2010). One possible reason for this is that some of the facial manipulations used to measure the aftereffects may also engage more generic levels of object processing.

Are there levels too high to account for the aftereffects? Fox and Barton (2007) explored expression aftereffects for a range of different stimuli. They found strong adaptation between faces of the same individual with different expressions, even when the adapt and test faces were drawn from different pictures of the person, again arguing against a purely low-level account of the adaptation. However, expression aftereffects in the faces did not occur when observers instead adapted to non-face stimuli with emotional meaning, suggesting that the biases were not generated at a conceptual level. On the other hand, other studies have found that gender aftereffects did transfer from headless male or female bodies to faces, implying adaptation to an abstract though still visual representation of gender (Ghuman, McDaniel, & Matin, 2010). Notably, even more abstract adaptation has been reported for motion—viewing still images that imply movement (Winawer, Huk, & Boroditsky, 2008), or even reading passages conveying movement (Dils & Boroditsky, 2010)—can induce a motion aftereffect. In summary, then, the site of the adaptation has been difficult to pin down, perhaps because adaptation itself is intrinsic to visual coding and thus may arise throughout the visual stream (Webster, 2011, 2015). In an early review of visual aftereffects, Mollon (1974) noted that the psychophysicist's obsession with adaptation stemmed from the belief that "if you can adapt it, it's there." As the horizon of possible aftereffects continues to expand, a possible corollary is that "if it's there, you can adapt it."

Recent studies have also explored the sites of face adaptation by assessing how it depends on the ability of the individual to recognize faces. Children with autism show impaired face recognition and may show weaker aftereffects (Pellicano, Jeffery, Burr, & Rhodes, 2007) that are specific to faces (Ewing, Pellicano, & Rhodes, 2013). These differences may also occur in their relatives (Fiorentini, Gray, Rhodes, Jeffery, & Pellicano, 2012). More extreme impairments occur in prosopagnosia, which is a disorder defined specifically by poor face recognition. Some patients may nevertheless experience face aftereffects comparable to normal subjects (Susilo et al., 2011), though for others these can differ in whether the faces are varied by distortions or actual changes in identity (Palermo, Rivolta, Wilson, & Jeffery, 2011). There are also large and reliable individual differences in face recognition among the normal population, and this capacity is weakly correlated with differences in the strength of eye-height aftereffects, which may strongly engage face-specific processing (Dennett, McKone, Edwards, & Susilo, 2012). It is difficult in all such experiments to rule out extraneous factors such as how and where the individual scans the face. Yet these studies raise the intriguing possibility that the proficiency and plasticity of face recognition may both depend on the form of the neural mechanisms underlying face coding.

Regardless of the basis of face aftereffects, the fact that adaptation can strongly bias the perception of faces—arguably among the most important social stimuli for humans—suggests that this adaptation may play an important role in how we normally experience and respond to faces. But are there signs that these adjustments occur outside the lab? A number of researchers have begun to explore face adaptation in more natural contexts (Carbon & Ditye 2012; Morikawa 2005; Webster et al., 2004). Adaptation predicts that the face we perceive as average should depend on the average face we have experienced, and this can account for individual differences in judgments of gender, attractiveness, and ethnicity (Apicella, Little, & Marlowe, 2007; Saxton, Little, DeBruine, Jones, & Roberts, 2009; Tovee, Swami, Furnham, & Mangalparsad, 2006; Webster et al., 2004). It may be callous to say that someone has "a face only a mother could love," but it is perhaps also insightful, for the illusion of beauty is that attractive faces are only average (Langlois & Roggman, 1990), and the illusion of adaptation is that what looks average is strongly weighted by the images we most often encounter.

REFERENCES

Afraz, A., & Cavanagh, P. (2009). The gender-specific face aftereffect is based in retinotopic not spatiotopic coordinates across several natural image transformations. *Journal of Vision, 9*(10):10, 11–17.

Apicella, C. L., Little, A. C., & Marlowe, F. W. (2007). Facial averageness and attractiveness in an isolated population of hunter-gatherers. *Perception, 36*(12), 1813–1820.

Bestelmeyer, P. E. G., Jones, B. C., Debruine, L. M., Little, A. C., Perrett, D. I., Schneider, A., … Conway, C. A. (2008). Sex-contingent face aftereffects depend on perceptual category rather than structural encoding. *Cognition, 107*(1), 353–365.

Carbon, C. C., & Ditye, T. (2012). Face adaptation effects show strong and long-lasting transfer from lab to more ecological contexts. *Frontiers in Psychology, 3*, 3.

Carbon, C. C., & Leder, H. (2006). Last but not least—The Mona Lisa effect: Is "our" Lisa fame or fake? *Perception, 35*(3), 411–414.

Clifford, C. W. G. (2002). Perceptual adaptation: Motion parallels orientation. *Trends in Cognitive Sciences, 6*(3), 136–143.

Dennett, H. W., McKone, E., Edwards, M., & Susilo, T. (2012). Face aftereffects predict individual differences in face recognition ability. *Psychological Science*, *23*(11), 1279–1287.

Dickinson, J. E., Almeida, R. A., Bell, J., & Badcock, D. R. (2010). Global shape aftereffects have a local substrate: A tilt aftereffect field. *Journal of Vision*, *10*(13), 5.

Dils, A. T., & Boroditsky, L. (2010). Visual motion aftereffect from understanding motion language. *Proceedings of the National Academy of Sciences USA, 107*(37), 16396–16400.

Ewing, L., Pellicano, E., & Rhodes, G. (2013). Atypical updating of face representations with experience in children with autism. *Developmental science*, *16*(1), 116–123.

Fiorentini, C., Gray, L., Rhodes, G., Jeffery, L., & Pellicano, E. (2012). Reduced face identity aftereffects in relatives of children with autism. *Neuropsychologia*, 50(12), 2926–2932.

Fox, C. J., & Barton, J. J. (2007). What is adapted in face adaptation? The neural representations of expression in the human visual system. *Brain Research, 1127*(1), 80–89.

Ghuman, A. S., McDaniel, J. R., & Matin, A. (2010). Face adaptation without a face. *Current Biology, 20*(1), 32–36.

Graham, N. V. (1989). *Visual pattern analyzers*. Oxford: Oxford University Press.

Hole, G. J. (2011). Identity-specific face adaptation effects: Evidence for abstractive face representations. *Cognition, 119*(2), 216–228.

Hsu, S. M., & Young, A. W. (2004). Adaptation effects in facial expression recognition. *Visual Cognition, 11*(7), 871–899.

Jenkins, R., Beaver, J. D., & Calder, A. J. (2006). I thought you were looking at me: Direction-specific aftereffects in gaze perception. *Psychological Science*, *17*(6), 506–513.

Kohn, A. (2007). Visual adaptation: Physiology, mechanisms, and functional benefits. *Journal of Neurophysiology, 97*(5), 3155–3164.

Langlois, J. H., & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science, 1*, 115–121.

Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience, 4*(1), 89–94.

Leopold, D. A., Rhodes, G., Müller, K. M., & Jeffery, L. (2005). The dynamics of visual adaptation to faces. *Proceedings of the Royal Society B, 272*(1566), 897–904.

MacLin, O. H., & Webster, M. A. (2001). Influence of adaptation on the perception of distortions in natural images. *Journal of Electronic Imaging, 10*(1), 100–109.

McKone, E., Jeffery, L., Boeing, A., Clifford, C. W., & Rhodes, G. (2014). Face identity aftereffects increase monotonically with adaptor extremity over, but not beyond, the range of natural faces. *Vision research*, *98*, 1–13.

Mollon, J. D. (1974). After-effects and the brain. *New Scientist, 61*, 479–482.

Moradi, F., Koch, C., & Shimojo, S. (2005). Face adaptation depends on seeing the face. *Neuron, 45*(1), 169–175.

Morikawa, K. (2005). Adaptation to asymmetrically distorted faces and its lack of effect on mirror images. *Vision Research, 45*(25–26), 3180–3188.

Müller, K. M., Schillinger, F., Do, D. H., & Leopold, D. A. (2009). Dissociable perceptual effects of visual adaptation. *PLoS ONE, 4*(7), e6183.

O'Neil, S., & Webster, M. A. (2011). Adaptation and the perception of facial age. *Visual Cognition, 19*(4), 534–550.

O'Neil, S. F., Mac, A., Rhodes, G., & Webster, M. A. (2014). Adding years to your life (or at least looking like it): a simple normalization underlies adaptation to facial age. *PloS One*, *9*(12), e116105.

O'Neil, S. F., Mac, A., Rhodes, G., & Webster, M. A. (2015). Model fitting versus curve fitting: A model of renormalization provides a better account of age aftereffects than a model of local repulsion. *i-Perception*, *6*(6), 2041669515613669.

Palermo, R., Rivolta, D., Wilson, C. E., & Jeffery, L. (2011). Adaptive face space coding in congenital prosopagnosia: Typical figural aftereffects but abnormal identity aftereffects. *Neuropsychologia, 49*(14), 3801–3812.

Pellicano, E., Jeffery, L., Burr, D., & Rhodes, G. (2007). Abnormal adaptive face-coding mechanisms in children with autism spectrum disorder. *Current Biology, 17*(17), 1508–1512.

Rhodes, G., Jeffery, L., Watson, T. L., Clifford, C. W., & Nakayama, K. (2003). Fitting the mind to the world: Face adaptation and attractiveness aftereffects. *Psychological Science, 14*(6), 558–566.

Rhodes, G., Robbins, R., Watson, T. L., Clifford, C. W., & Nakayama, K. (2005). Adaptation and face perception: How aftereffects implicate norm based coding of faces. In C. W. G. Clifford & G. Rhodes (Eds.), *Fitting the mind to the world: Adaptation and aftereffects in high-level vision* (pp. 213–240). Oxford: Oxford University Press.

Robbins, R., McKone, E., & Edwards, M. (2007). Aftereffects for face attributes with different natural variability: Adapter position effects and neural models. *Journal of Experimental Psychology: Human Perception and Performance, 33*(3), 570–592.

Ross, D. A., Deroche, M., & Palmeri, T. J. (2014). Not just the norm: Exemplar-based models also predict face aftereffects. *Psychonomic bulletin & review*, *21*(1), 47-70.

Saxton, T. K., Little, A. C., DeBruine, L. M., Jones, B. C., & Roberts, C. (2009). Adolescents' preferences for sexual dimorphism are influenced by relative exposure to male and female faces. *Personality and Individual Differences, 47*(8), 864–868.

Schwartz, O., Hsu, A., & Dayan, P. (2007). Space and time in visual context. *Nature Reviews Neuroscience, 8*(7), 522–535.

Schweinberger, S. R., Zaske, R., Walther, C., Golle, J., Kovács, G., & Wiese, H. (2010). Young without plastic surgery: Perceptual adaptation to the age of female and male faces. *Vision Research, 50*(23), 2570–2576.

Simas, M. L. (2000). The multiple-faces phenomenon: Some investigate studies. *Perception, 29*(11), 1393–1395.

Storrs, K. R., & Arnold, D. H. (2012). Not all face aftereffects are equal. *Vision Research, 64*, 7–16.

Strobach, T., & Carbon, C. C. (2013). Face adaptation effects: reviewing the impact of adapting information, time, and transfer. *Frontiers in psychology*,*4*.

Susilo, T., McKone, E., Dennett, H., Darke, H., Palermo, R., Hall, A., Rhodes, G. (2011). Face recognition impairments despite normal holistic processing and face space coding: Evidence from a case of developmental prosopagnosia. *Cognitive Neuropsychology, 27*(8), 636–664.

Susilo, T., McKone, E., & Edwards, M. (2010). Solving the upside-down puzzle: Why do upright and inverted face aftereffects look alike? *Journal of Vision*, *10*(13), 1.

Tangen, J. M., Murphy, S. C., & Thompson, M. B. (2011). Flashed face distortion effect: Grotesque faces from relative spaces. *Perception, 40*(5), 628–630.

Tovee, M. J., Swami, V., Furnham, A., & Mangalparsad, R. (2006). Changing perceptions of attractiveness as observers are exposed to a different culture. *Evolution and Human Behavior, 27*, 443–456.

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Quarterly Journal of Experimental Psychology A, 43*(2), 161–204.

Valentine, T., Lewis, M. B., & Hills, P. J. (2015). Face-space: A unifying concept in face recognition research. *The Quarterly Journal of Experimental Psychology*, (ahead-of-print), 1-24.

Watson, T. L., & Clifford, C. W. G. (2003). Pulling faces: An investigation of the face-distortion aftereffect. *Perception, 32*(9), 1109–1116.

Webster, M. A. (2011). Adaptation and visual coding. *Journal of Vision*, *11*(5):3, 1–23.

Webster, M. A. (2015). Visual Adaptation. Annual Review of Vision Science1, 547-567.

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature, 428*(6982), 557–561.

Webster, M. A., & MacLeod, D. I. A. (2011). Visual adaptation and face perception. *Philosophical Transactions of the Royal Society B*, *366*(1571), 1702–1725.

Webster, M. A., & MacLin, O. H. (1999). Figural aftereffects in the perception of faces. *Psychonomic Bulletin and Review, 6*(4), 647–653.

Winawer, J., Huk, A. C., & Boroditsky, L. (2008). A motion aftereffect from still photographs depicting motion. *Psychological Science, 19*(3), 276–283.

Xu, H., Dayan, P., Lipkin, R. M., & Qian, N. (2008). Adaptation across the cortical hierarchy: Low-level curve adaptation affects high-level facial-expression judgments. *Journal of Neuroscience, 28*(13), 3374–3383.

Zhao, C., Series, P., Hancock, P. J., & Bednar, J. A. (2011). Similar neural adaptation mechanisms underlying face gender and tilt aftereffects. *Vision Research, 51*(18), 2021–2030.

Zhao, L., & Chubb, C. (2001). The size-tuning of the face-distortion after-effect. *Vision Research, 41*(23), 2979–2994.

PART VI
Grouping and Organization

Chapter 95
Ambiguous Figures Moving Forward

Lori J. Bernstein

EFFECT OF MOTION ON AMBIGUOUS FIGURES

Visual perceptual identification is a dynamic process that involves analyzing and synthesizing elementary features of sensory information to yield a cohesive percept of the external world. Our perceptions, recognitions, and identifications of the objects we see are usually based on shape, but those processes also can be influenced by other types of information. Direction of motion is one example of non-shape information that can influence the perceptual interpretation, or "construal" of an object, and this is most noticeable when the identity of the object is ambiguous.

The role of directional motion in the identification of ambiguous figures was investigated in a series of experiments in which participants were shown figures of ambiguous identity moving in different directions along the horizontal plane (Bernstein & Cooper, 1997). The first experiment used a long-range induced motion paradigm in which a square frame was displaced around a figure (Duncker, 1938). Two discrete presentations of the frame appeared in different positions with respect to the figure, which itself did not move. This produced induced motion of the figure: although the figure did not change positions, it was perceived as having moved (in the opposite direction of the frame). Under these conditions, more often than not, when a figure is perceived as moving to the right, it is identified as the object whose face is on the right; when the display induces leftward motion, the percept is that object that faces to the left. The observer construes the figure to be the one that is heading in the direction of motion; hence the phenomenon was called the *heading* effect (Bernstein & Cooper, 1997). The heading effect also holds under conditions when the figure changes positions and there is no frame inducing the perception of motion.

Six ambiguous two-dimensional line drawings were used in the original heading effect studies (see Fig. VI.95-1). The goose-hawk is from a study with goslings (Tinbergen, 1939, 1951), which is explained in more detail later in this chapter. The duck-rabbit (McBeath, Morikawa, & Kaiser, 1992) was simplified from the earliest known duck-rabbit drawing appearing in a German humor magazine (Fliegende Blätter, 1892) but usually attributed to Jastrow (1899). The face-body was modified from a card published by Y & B Associates in 1987, although the original source has not been identified. Bernstein and Cooper (1997) created three original ambiguous figures (a "cat-swan," "snail-whale," and "ice cream truck–semi trailer truck"). The heading effect generalized to all these ambiguous figures that had left and right directionality, although the strength of the bias varied. Interactive Figure VI.95-1 shows ambiguous figures moving from right to left or from left to right across the screen.

RELEVANCE OF THE HEADING EFFECT FOR VISION RESEARCH

We have long known that motion can influence object perception. For instance, kinetic depth and "structure-through-motion" experiments show that the visual system

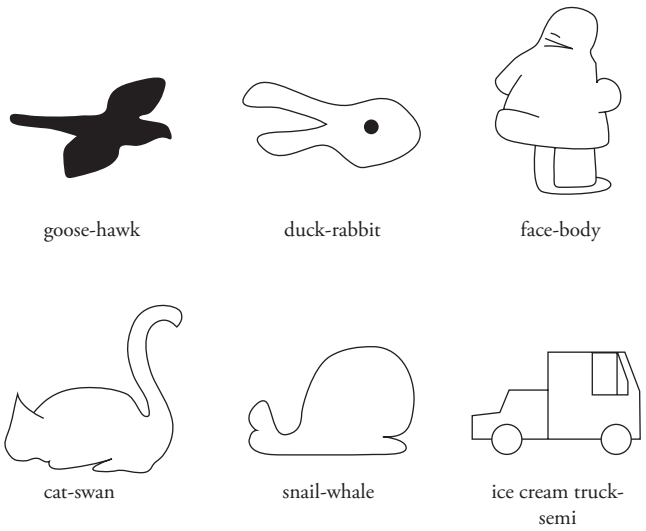


Figure VI.95-1. In all cases, the direction of the ambiguous figure's perceived translational motion increases the probability that observers construe the figure as that object whose front faces in the portrayed direction of motion ("heading effect"). If perceived figural motion is right, participants are more likely to see the figures as: hawk, rabbit, body, swan, whale, and semi. If perceived figural motion is left, the figures are more likely to be perceived as: goose, duck, face, cat, snail, and ice cream truck. Vertical motion has no consistent impact on perceptual identification of these figures. (Figure by Bernstein, L. J., & Cooper, L. A. (1997). Direction of motion influences perceptual identification of ambiguous figures. *Journal of Experimental Psychology: Human Perception and Performance, 23*(3), 721–737.)



Interactive Figure VI.95-1. Heading Effect Demonstration: Ambiguous figure is more likely to be construed as the object heading in the direction of motion.

perceives a two-dimensional form when that object is static but a three-dimensional form when the object is moving (i.e., Cutting, 1978; Hill & Johnston, 2001; Johansson, 1973; Ullman, 1979; Wallach & O'Connell, 1953). However, unlike the right or left translational motion that elicits the heading effect, these other types of motions provide new information such as different viewpoints of the objects through changing rotation in depth and/or new geometric relationships among elements. The heading effect shows that even relatively simple direction of motion can impact something as complex as shape perception. Interestingly, the reciprocal occurs; that is, the shape of a figure can impact the perception of direction of translational motion when the direction of that motion is ambiguous and when the object's identity is associated with a probable path of motion (e.g., Brosgole & Whalen, 1967; Jones & Bruner, 1954; McBeath et al., 1992; Shiffrar & Freyd, 1993).

In most situations, the shape of an object alone is sufficient to identify objects correctly. Theories of visual object recognition and perception reflect this fact as they try to explain how the perceptual system can derive a mental representation of three-dimensional objects from two-dimensional retinal projections. This is true for theories that argue that the information relevant for recognition is derived from the structure of objects or the relationship among axes of elongation, symmetry, and object parts (e.g., Bucher & Palmer, 1985; Marr & Nishihara, 1978; Palmer, 1985). This is also true for theories that emphasize the role of parts themselves and local relations among parts (e.g., Biederman, 1987). Still, objects have properties other than shape (e.g., color, size, texture, and orientation), and considerable evidence demonstrates that such factors influence object perception as well as memory of them (e.g., Cooper, Schacter, Ballesteros, & Moore, 1992; Garner, 1978; Jolicoeur, 1987; Knappmeyer, Thornton, & Bulthoff, 2003; Tarr & Pinker, 1989; Treisman & Souther, 1986; Wohlschlager, 2000). A wide variety of research approaches and techniques have focused on elucidating how the visual system analyzes and synthesizes input. The research has focused on the mechanisms by which different sources of information, including but not limited to shape and motion, are combined to yield accurate mental representations of the world, including animal models, computational modeling, behavioral, and/or neuroimaging techniques in healthy people and those who have had a stroke.

VARIANTS AND PARAMETERS

Construing ambiguous figures as the object that is heading in the perceived motion direction is relatively robust across a variety of motion paradigms, stimuli, and task procedures. It has been observed under conditions of long-range apparent motion, induced motion, and continuous motion (Bernstein & Cooper, 1997). The induced motion paradigm shows that the ambiguous figure need not change position to produce the heading effect, but perceiving that the object has moved appears necessary. There are temporal limitations to the heading effect when the motion display uses an apparent motion paradigm. As the interstimulus-interval (ISI) increases, the effect dissipates, and one experiment showed the heading effect disappearing somewhere between 1,000 and 1,500 ms (Bernstein & Cooper, 1997), far longer than predicted if compelling motion is required.

In order to further push the generalizability of the effect, four new snail-whale figures on a continuum ranging from being an unambiguous snail to being an unambiguous whale were created (Bernstein & Cooper, 1997). While direction of motion had no impact on how the figure of an unambiguous snail or whale was identified, motion direction did have an effect on interpretation even when the form partially biased one interpretation over the other. That is, motion influences form interpretation, even when some "what" information is available. More recent studies using other types of ambiguous figures have found results consistent with this (e.g., Newell, Wallraven, & Huber, 2004). In the original Bernstein and Cooper (1997) studies, the ambiguous figures presented had two alternatives in terms of identification, and participants in those studies were explicitly made aware of both of them in advance. Since those initial studies, it has been shown that the heading effect persists when there is no obvious forced-choice alternative of identification and when observers are naïve to the stimuli. For instance, both adults and young children identify impoverished point-light biological motion stimuli as objects facing in the direction of apparent motion (Pavlova, Krageloh-Mann, Birbaumer, & Sokolov, 2002) and when there is no a priori association of an object with a particular motion, such as an equilateral triangle (Bucher & Palmer, 1985). In addition, recognition of objects is hampered when those objects are shown moving in a direction inconsistent with a forward direction (Verfaillie, 2000).

The heading effect emerges when the task involves simple translational motion of familiar bistable figures that face opposite directions. Forced-choice object construal is likely not identical to the processes of identifying or recognizing a shape when the alternatives are not known and when there is no obvious alternative identity associated with a shape. Nonetheless, the heading effect remains relatively robust across a range of shape stimuli, temporal motion conditions, and people of different ages and experience.

OTHER CONSIDERATIONS

The first evidence that motion and shape pairings affect object perception came from animal-behavior studies. Tingbergen (1939, 1951) found that the silhouette of an ambiguous goose-hawk figure (shown in Fig. VI.95-1) elicited a fear response in goslings when moved overhead in the shorter-part-first direction, but it did not elicit fear when moved in the longer-part-first direction. The short neck and long tail are characteristic of a predator (e.g., a hawk),

whereas the long-necked and short-tailed figure is more characteristic of a nonpredator (e.g., a goose). It was not known whether the goslings' fear response occurred because they identified the object as a predator or because the particular motion-shape combination was novel. It was clear, however, that the baby geese discriminated on the basis of shape-motion contingencies (i.e., the same shape eliciting different reaction depending on its direction of motion).

The heading effect demonstrates that direction of motion is a stimulus factor that biases the identification of ambiguous figures, although it does not explain the cognitive mechanisms by which this occurs. Some possibilities include top-down inference (e.g., construal is based on knowledge that objects usually move in a forward-facing direction) and visual attention (e.g., construal is driven by where attention is focused with respect to the figure; Bernstein & Cooper, 1997). Some of the displays in which the heading effect remains effective fall outside traditional long-range ISI conditions in which apparent motion is seen, but signals that keep attention positioned over an object provide information about displacement or location, which may then produce an impression of motion (see Cavanagh, 1992; Lu & Sperling, 1995). The healthy visual system is flexible enough to allow considerable time spread and interruption between two presentations to perceive motion, yet there are temporal limits to this flexibility. Exceeding some temporal limit may result in successive presentations no longer being seen either as one object moving or the same object whose position has changed. In the absence of motion, interpretation of an object that has ambiguous or inconsistent local information depends on the part of that object the viewer elects to focus on or is instructed to attend (Peterson & Hochberg, 1983; Tsal & Kolbet, 1985; Xu & Franconeri, 2012). It has been suggested that stimulus motion works primarily as a categorical, directional expectation, which may direct visual attention (Hochberg, 2003). In any case, the information and processing that enables the heading effect diminish as ISI increases.

How the shape and motion integration occurs and at what level in the visual system is also equivocal. More basic neuroscience studies support the independence of motion and shape processing. For instance, some visual cortex neurons respond identically to motion across their receptive fields, regardless of shape (Hubel & Wiesel, 1959, 1962, 1968), whereas other neurons selectively respond to particular shapes, regardless of motion (Schiller, Finlay, & Volman, 1976a, 1976b, 1976c). The brain appears to have distinct pathways ("what versus where") in function and location. The magnocellular pathway is suited more to motion detection and is more ventral, whereas the parvocellular pathway is more sensitive to form information and is located more dorsally (Benardete, Kaplan, & Knight, 1992; Lee, Pokorny, Smith, & Kremers, 1994). Positron emission tomography and functional magnetic resonance imaging have shown that different brain areas become active as participants are required to respond to different aspects (e.g., shape, color, or motion) of the same display (e.g., Corbetta, Miezin, Shulman, & Petersen, 1991; Gulyas & Roland, 1991). Areas of activity observed in humans correspond to those found in animal studies.

Neuropsychology case studies provide compelling data that different features of objects are processed by modular brain systems that are largely independent. For instance, people with focal cortical damage following stroke may present with selective deficits of processing individual features, including color (achromatopsia; Bouvier & Engel, 2006; Damasio, Yamada, Damasio, Corbett, & McKee, 1980; Meadows, 1974; Zeki, 1990) or form (visual form agnosia; Goodale, Milner, Jakobson, & Carey, 1991; Karnath, Ruter, Mandler, & Himmelbach, 2009; Riddoch & Humphreys, 1987; Riddoch et al., 2008). Motion detection can be selectively impaired following bilateral damage to the MT human homologue (occipital-temporal lobe), with intact perception of other features (akinetopsia; McLeod, Heywood, Driver, & Zihl, 1989; Vaina, 1989; Zihl, von Cramon, & Mai, 1983; Zihl, von Cramon, Mai, & Schmid, 1991). In one case of akinetopsia (Zihl et al., 1983), the person reported perceiving running water as a static icicle.

On the other hand, a growing body of evidence from behavioral (Lorenceau & Alais, 2001; Stone & Harper, 1999; Wallis & Bulthoff, 2001), computational modeling (Giese & Poggio, 2003), and neural studies (Bradley, Chang, & Andersen, 1998; Decety & Grezes, 1999; Haxby, Petit, Ungerleider, & Courtney, 2000; Kourtzi, Bulthoff, Erb, & Grodd, 2002; Oram & Perrett, 1996) is more consistent with form and motion information combining when recognizing different types of objects. For instance, studies examining identification of faces from either facial-form features and/or facial motion have shown that when both types of information are available, responses reflect input from both types of information and represent some compromise (Knappmeyer et al., 2003). Together, the literature indicates that form and motion interact in numerous ways at different levels in the visual system, with motion impacting shape perception and shape impacting motion perception.

On balance, findings support modularity of feature representations in the human brain until the parietal lobes and their spatial and/or attentional representations permit accurate binding of those different features (Friedman-Hill, Robertson, & Treisman, 1995; Shafritz, Gore, & Marois, 2002; Treisman & Gelade, 1980). Bernstein and Robertson (1998) presented a case study of a patient (RM) with intact primary-vision despite bilateral parietal-occipital lesions. RM was presented with two items that moved in different ways at the same time. He often miscombined the shape, color, and/or motion of items, seeing the shape and/or the color of one of the items with the other item's motion. His pattern of errors, known as illusory conjunctions, showed that correctly combining two features did not necessarily lead to correctly binding the third. In this case, the three features remained independent. One may speculate that intact parietal lobes are necessary to support the heading effect.

In healthy people, we know that information is integrated effortlessly. The work with RM suggests that parietal lobes are required for integration of shape and motion because this integration is accomplished through an accurate spatial representation (Robertson, 2003). Thus higher cortical areas, whether human or aviary (i.e., Tinbergen's geese), are implicated in the integration of motion and shape information. Recent work with zebrafish, however, calls this into question. Fish (who have no cortical brain structures) preferred images of other zebrafish only when the form and motion were presented together in a meaningful manner (Neri, 2012).

In summary, there is a growing body of evidence that form and motion information interact with each other in complex ways and across multiple levels of visual processing. There are several issues that deserve further study. The levels at which form and motion interact remain to be determined. Research using neuroimaging techniques may shed light on the neural architecture that underlies the processing and integration of form and motion information. Ambiguous figures remind us that perception of objects involves much more than sensation because the information available to the eye is identical, yet the perception is not. We cannot ask cells or animals to report on what they "see," and using ambiguous stimuli is a clever way to probe these issues. Little is known about the development of these processes and how they change with normal and/or pathological development and aging. For instance, one might predict that older adults would be less sensitive to the heading effect because of age-related declines in motion perception (e.g., Bennett, Sekuler, & Sekuler, 2007; Snowden & Kavanagh, 2006), including reduced spatiotemporal range yielding good direction discrimination (Roudaia, Bennett, Sekuler, & Pilz, 2010) or changes in visual spatial attention (e.g., Greenwood, Parasuraman, & Alexander, 1997; Plude & Doussard-Roosevelt, 1989; Trick & Ens, 1998). Finally, paradigms and figures of varying levels of ambiguity described here and elsewhere (e.g., Stottinger et al., 2014) may provide a foundation to develop novel ways of assessing complex visual processes in healthy and patient populations that can yield new information not afforded by traditional psychometric or neuropsychological tests.

REFERENCES

Benardete, E. A., Kaplan, E., & Knight, B. W. (1992). Contrast gain control in the primate retina: P cells are not X-like, some M cells are. *Visual Neuroscience*, *8*(5), 483–486.

Bennett, P. J., Sekuler, R., & Sekuler, A. B. (2007). The effects of aging on motion detection and direction identification. *Vision Research*, *47*(6), 799–809.

Bernstein, L. J., & Cooper, L. A. (1997). Direction of motion influences perceptual identification of ambiguous figures. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(3), 721–737.

Bernstein, L. J., & Robertson, L. C. (1998). Illusory conjunctions of color and motion with shape following bilateral parietal lesions. *Psychological Science*, *9*(3), 167–175.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*(2), 115–147.

Bouvier, S. E., & Engel, S. A. (2006). Behavioral deficits and cortical damage loci in cerebral achromatopsia. *Cerebral Cortex*, *16*(2), 183–191.

Bradley, D. C., Chang, G. C., & Andersen, R. A. (1998). Encoding of three-dimensional structure-from-motion by primate area MT neurons. *Nature*, *392*(6677), 714–717.

Brosgole, L., & Whalen, P. M. (1967). The effect of meaning on the allocation of visually induced motion. *Perception & Psychophysics*, *2*, 275–277.

Bucher, N. M., & Palmer, S. E. (1985). Effects of motion on perceived pointing of ambiguous triangles. *Perception & Psychophysics*, *38*(3), 227–236.

Cavanagh, P. (1992). Attention-based motion perception. *Science*, *257*(5076), 1563–1565.

Cooper, L. A., Schacter, D. L., Ballesteros, S., & Moore, C. (1992). Priming and recognition of transformed three-dimensional objects: Effects of size and reflection. *Journal of Experimental Psychology. Learning, Memory, and Cognition.18*(1), 43–57.

Corbetta, M., Miezin, F. M., Shulman, G. L., & Petersen, S. E. (1991). Selective attention modulates extrastriate visual regions in humans during visual feature discrimination and recognition. *Ciba Foundation Symposium*, *163*, 165–175; discussion 175–180.

Cutting, J. E. (1978). Generation of synthetic male and female walkers through manipulation of a biomechanical invariant. *Perception*, *7*(4), 393–405.

Damasio, A., Yamada, T., Damasio, H., Corbett, J., & McKee, J. (1980). Central achromatopsia: Behavioral, anatomic, and physiologic aspects. *Neurology*, *30*(10), 1064–1071.

Decety, J., & Grezes, J. (1999). Neural mechanisms subserving the perception of human actions. *Trends in Cognitive Science*, *3*(5), 172–178.

Duncker, K. (1938). Induced motion. In W. D. Ellis (Ed.), *A source book of Gestalt psychology* (pp. 161–172). New York: Harcourt, Brace.

Fliegende Blätter. (1892). *Welche Thiere gleichen einander am meisten? ("Which animals are most like each other?").* 23 October 1892, Munich, Germany

Friedman-Hill, S. R., Robertson, L. C., & Treisman, A. (1995). Parietal contributions to visual feature binding: Evidence from a patient with bilateral lesions. *Science*, *269*(5225), 853–855.

Garner, W. R. (1978). Selective attention to attributes and to stimuli. *Journal of Experimental Psychology: General*, *107*(3), 287–308.

Giese, M. A., & Poggio, T. (2003). Neural mechanisms for the recognition of biological movements. *Nature Reviews Neuroscience*, *4*(3), 179–192.

Goodale, M. A., Milner, A. D., Jakobson, L. S., & Carey, D. P. (1991). A neurological dissociation between perceiving objects and grasping them. *Nature*, *349*(6305), 154–156.

Greenwood, P. M., Parasuraman, R., & Alexander, G. E. (1997). Controlling the focus of spatial attention during visual search: Effects of advanced aging and Alzheimer disease. *Neuropsychology*, *11*(1), 3–12.

Gulyas, B., & Roland, P. E. (1991). Cortical fields participating in form and colour discrimination in the human brain. *Neuroreport*, *2*(10), 585–588.

Haxby, J. V., Petit, L., Ungerleider, L. G., & Courtney, S. M. (2000). Distinguishing the functional roles of multiple regions in distributed neural systems for visual working memory. *NeuroImage*, *11*(2), 145–156.

Hill, H., & Johnston, A. (2001). Categorizing sex and identity from the biological motion of faces. *Current Biology*, *11*(11), 880–885.

Hochberg, J. (2003). Acts of perceptual inquiry: Problems for any stimulus-based simplicity theory. *Acta Psychologica*, *114*(3), 215–228.

Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology*, *148*, 574–591.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, *160*, 106–154.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, *195*(1), 215–243.

Jastrow, J. (1899). The mind's eye. *Popular Science Monthly*, *54*, 299–312.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, *14*, 201–211.

Jolicoeur, P. (1987). A size-congruency effect in memory for visual shape. *Memory and Cognition*, *15*, 531–543.

Jones, E. E., & Bruner, J. S. (1954). Expectancy in apparent visual movement. *British Journal of Psychology*, *45*(3), 157–165.

Karnath, H. O., Ruter, J., Mandler, A., & Himmelbach, M. (2009). The anatomy of object recognition—Visual form agnosia caused by medial occipitotemporal stroke. *Journal of Neuroscience*, *29*(18), 5854–5862.

Knappmeyer, B., Thornton, I. M., & Bulthoff, H. H. (2003). The use of facial motion and facial form during the processing of identity. *Vision Research*, *43*(18), 1921–1936.

Kourtzi, Z., Bulthoff, H. H., Erb, M., & Grodd, W. (2002). Object-selective responses in the human motion area MT/MST. *Nature Neuroscience*, *5*(1), 17–18.

Lee, B. B., Pokorny, J., Smith, V. C., & Kremers, J. (1994). Responses to pulses and sinusoids in macaque ganglion cells. *Vision Research*, *34*(23), 3081–3096.

Lorenceau, J., & Alais, D. (2001). Form constraints in motion binding. *Nature Neuroscience*, *4*(7), 745–751.

Lu, Z. L., & Sperling, G. (1995). Attention-generated apparent motion. *Nature*, *377*(6546), 237–239.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London, B*, *200*(1140), 269–294.

McBeath, M. K., Morikawa, K., & Kaiser, M. K. (1992). Perceptual bias for forward-facing motion. *Psychological Science*, *3*(6), 362–367.

McLeod, P., Heywood, C., Driver, J., & Zihl, J. (1989). Selective deficit of visual search in moving displays after extrastriate damage. *Nature*, *339*(6224), 466–467.

Meadows, J. C. (1974). Disturbed perception of colours associated with localized cerebral lesions. *Brain*, *97*(4), 615–632.

Neri, P. (2012). Feature binding in zebrafish. *Animal Behavior*, *84*, 485–493.

Newell, F. N., Wallraven, C., & Huber, S. (2004). The role of characteristic motion in object categorization. *Journal of Vision*, *4*(2), 118–129.

Oram, M. W., & Perrett, D. I. (1996). Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. *Journal of Neurophysiology*, *76*(1), 109–129.

Palmer, S. E. (1985). The role of symmetry in shape perception. *Acta Psychologica*, *59*(1), 67–90.

Pavlova, M., Krageloh-Mann, I., Birbaumer, N., & Sokolov, A. (2002). Biological motion shown backwards: The apparent-facing effect. *Perception*, *31*(4), 435–443.

Peterson, M. A., & Hochberg, J. (1983). Opposed-set measurement procedure: A quantitative analysis of the role of local cues and intention in form perception. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 183–193.

Plude, D. J., & Doussard-Roosevelt, J. A. (1989). Aging, selective attention, and feature integration. *Psychology and Aging*, *4*(1), 98–105.

Riddoch, M. J., & Humphreys, G. W. (1987). A case of integrative visual agnosia. *Brain*, *110*(Pt. 6), 1431–1462.

Riddoch, M. J., Humphreys, G. W., Akhtar, N., Allen, H., Bracewell, R. M., & Schofield, A. J. (2008). A tale of two agnosias: Distinctions between form and integrative agnosia. *Cognitive Neuropsychology*, *25*(1), 56–92.

Robertson, L.C. (2003). Binding, spatial attention and perceptual awareness. *Nature Reviews Neuroscience*, *4*(2), 93–102

Roudaia, E., Bennett, P. J., Sekuler, A. B., & Pilz, K. S. (2010). Spatiotemporal properties of apparent motion perception and aging. *Journal of Vision*, *10*(14), 5.

Schiller, P. H., Finlay, B. L., & Volman, S. F. (1976a). Quantitative studies of single-cell properties in monkey striate cortex: I. Spatiotemporal organization of receptive fields. *Journal of Neurophysiology*, *39*(6), 128–1319.

Schiller, P. H., Finlay, B. L., & Volman, S. F. (1976b). Quantitative studies of single-cell properties in monkey striate cortex: II. Orientation specificity and ocular dominance. *Journal of Neurophysiology*, *39*(6), 132–1333.

Schiller, P. H., Finlay, B. L., & Volman, S. F. (1976c). Quantitative studies of single-cell properties in monkey striate cortex: III. Spatial frequency. *Journal of Neurophysiology*, *39*(6), 133–1351.

Shafritz, K. M., Gore, J. C., & Marois, R. (2002). The role of the parietal cortex in visual feature binding. *Proceedings of the National Academy of Sciences USA*, *99*(16), 1091–10922.

Shiffrar, M., & Freyd, J. J. (1993). Timing and apparent motion path choice with human body photographs. *Psychological Science*, *4*, 37–384.

Snowden, R. J., & Kavanagh, E. (2006). Motion perception in the ageing visual system: Minimum motion, motion coherence, and speed discrimination thresholds. *Perception*, *35*(1), 9–24.

Stone, J. V., & Harper, N. (1999). Temporal constraints on visual learning: A computational model. *Perception*, *28*(9), 1089–1104.

Stottinger, E., Filipowicz, A., Marandi, E., Quehl, N., Danckert, J., & Anderson, B. (2014). Statistical and perceptual updating: Correlated impairments in right brain injury. *Experimental Brain Research*, *232*(6), 1971–1978.

Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology*, *21*(2), 233–282.

Tinbergen, N. (1939). Why do birds behave the way they do? *Bird Lore*, *41*, 23–30.

Tinbergen, N. (1951). *The study of instict*. Oxford: Oxford University Press.

Treisman, A., & Souther, J. (1986). Illusory words: The roles of attention and of top-down constraints in conjoining letters to form words. *Journal of Experimental Psychology: Human Perception and Performance*, *12*(1), 3–17.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136.

Trick, L. M., & Ens, J. T. (1998). Lifespan changes in attention: The visual search task. *Cognitive Development*, *13*, 369–386.

Tsal, Y., & Kolbet, L. (1985). Disambiguating ambiguous figures by selective attention. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *37*(1), 25–37.

Ullman, S. (1979). The interpretation of structure from motion. *Proceedings of the Royal Society of London, B*, *203*(1153), 405–426.

Vaina, L. M. (1989). Selective impairment of visual motion interpretation following lesions of the right occipito-parietal area in humans. *Biological Cybernetics*, *61*(5), 347–359.

Verfaillie, K. (2000). Perceiving human locomotion: Priming effects in direction discrimination. *Brain Cognition*, *44*(2), 192–213.

Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, *45*(4), 205–217.

Wallis, G., & Bulthoff, H. H. (2001). Effects of temporal association on recognition memory. *Proceedings of the National Academy of Sciences USA*, *98*(8), 4800–4804.

Wohlschlager, A. (2000). Visual motion priming by invisible actions. *Vision Research*, *40*(8), 925–930.

Xu, Y., & Franconeri, S. L. (2012). The head of the table: Marking the "front" of an object is tightly linked with selection. *Journal of Neuroscience*, *32*(4), 1408–1412.

Zeki, S. (1990). A century of cerebral achromatopsia. *Brain*, *113*(Pt. 6), 1721–1777.

Zihl, J., von Cramon, D., & Mai, N. (1983). Selective disturbance of movement vision after bilateral brain damage. *Brain*, *106*(Pt. 2), 313–340.

Zihl, J., von Cramon, D., Mai, N., & Schmid, C. (1991). Disturbance of movement vision after bilateral posterior brain damage: Further evidence and follow up observations. *Brain*, *114*(Pt. 5), 2235–2252.

Chapter 96
The Scramble Illusion

Texture Metamers

Charles Chubb, Joseph Darcy, Michael S. Landy, John Econopouly, Jong-Ho Nam, Dan Bindman, and George Sperling

Challenge: Take a look at Figure VI.96-1. Can you explain what makes this figure interesting? This is a question that we have posed to many audiences since we first discovered the scramble illusion. No one has ever come up with the right answer.

As we shall show, the phenomenology of visual texture perception is strangely misleading. Wherever we look within a field of texture, we can see all the details of the texture filling that region. For example, when we look at the texture shown in Figure VI.96-1, we can see the specific gray levels of the squares in whichever region we choose to inspect. This gives us the impression that when we view a texture, our eyes and brain take a "visual inventory" of the texture and that this inventory defines the impact exerted on our vision by the texture. Under this naïve theory, two textures should look different if they differ significantly in the relative proportions of the various features they contain.

That the visual-inventory of features theory is false is dramatized by Figure VI.96-2. The left and right sides of Figure VI.96-2 are filled with two different types of texture called "scrambles." A scramble is composed of small squares of different gray levels mixed together in specified proportions and randomly arranged in space. The scrambles shown in Figure VI.96-2 appear to be homogeneous. However, careful scrutiny reveals that the scrambles on the left and right sides differ dramatically in ways that human viewers utterly fail to notice without explicit prompting. The actual histograms of the left and right sides in Figure VI.96-2 are shown in the inset bar graphs. The histograms give the proportions of different gray levels in the two scrambles. As shown by the right-hand bar graph, the scramble filling the right half of Figure VI.96-2 contains equal proportions of 17 different gray levels that increase from black to white. By contrast, the left-hand bar graph shows that the scramble filling the left side of this figure consists of a mixture of just three gray levels whose proportions have been carefully balanced to achieve several conditions that are described later.

Now that we have seen the two different scrambles in Figure VI.96-2, take another look at Figure VI.96-1. This figure also comprises two different scrambles of the same sorts as are used in Figure VI.96-2. Can you find the boundary between them?

THE SCRAMBLE ILLUSION AND PREATTENTIVE VISION

Figures VI.96-1 and VI.96-2 tell us a great deal about human preattentive vision. The word "preattentive" refers to those processes that operate, prior to any effort of attention on the part of the viewer, to make manifest the qualitative differences in the visual input that exist across space. White paint on a dark red background produces a visible difference that we cannot fail to see. Any difference of this sort that our vision reveals to us immediately and spontaneously is considered to be preattentive. It is clear that human preattentive vision is not at all sensitive to any of the physical differences that exist between the scrambles on the left-hand versus the right-hand side of Figure VI.96-2. We assume that in order for human vision to be preattentively sensitive to some particular physical property, there must exist a class of neurons operating in human preattentive vision that are differentially activated by variations in this statistic. In other words, any difference that is immediately evident to human vision implies the existence of some population of neurons sensitive to that difference. We take as a working hypothesis that preattentive human vision is devoid of any neurons sensitive to any of the many physical differences that exist between the right and left sides of Figure VI.96-2.

For example, we can immediately rule out the possibility that human preattentive vision contains a class of neurons that is specifically sensitive to any of the three gray levels that are used to make up the scramble on the left side of Figure VI.96-2. Suppose, in particular, that human preattentive vision contained a retinotopically organized array of neurons each of which was highly activated by gray level #15 in our set (third from right) and not at all by any other gray levels. The output from such an array would compose a "neural image" (Robson, 1980) providing

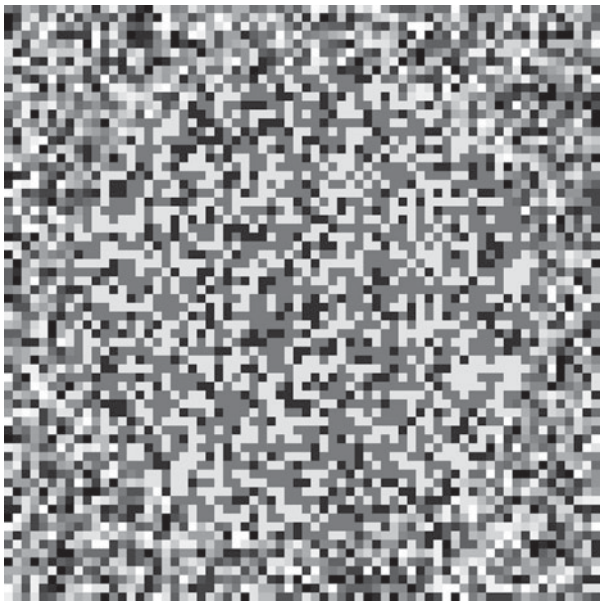


Figure VI.96-1. Can you guess what makes this figure interesting?

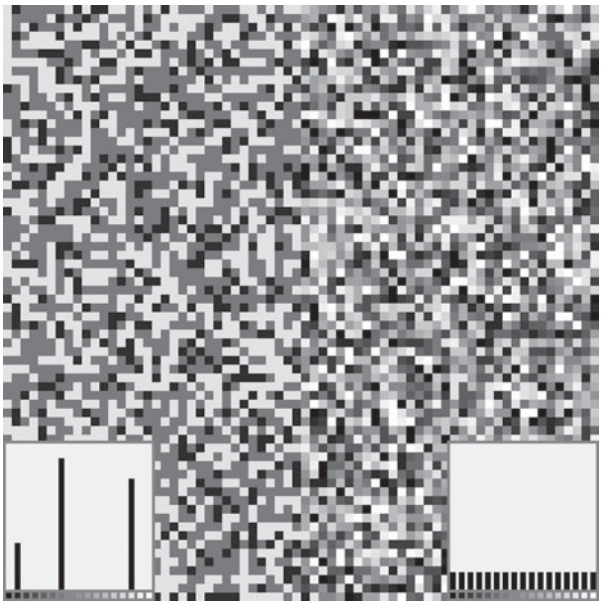


Figure VI.96-2. Two adjacent scrambles separated by a vertical boundary. The histogram insert at the lower left indicates the relative proportions of the three gray levels in the left half of the figure. The histogram insert at the lower right indicates that all 17 gray levels in the right half of the scramble have equal proportions.

a spatial map of gray level #15. Using black to code no activation and white to code the activation produced by gray level #15, the neural image produced in this array by Figure VI.96-2 would look like Figure VI.96-3. The large difference in average activation between the left and right sides of Figure VI.96-3 indicates that if human vision did possess such a neural array tuned specifically to gray level #15, then we would immediately see the difference between the left and right sides of Figure VI.96-2. We conclude that human preattentive vision contains no such neural array.



Figure VI.96-3. The scrambles of Figure VI.96-2 with gray level 15 (almost white) represented as white and all other gray levels represented as black.

Figures VI.96-1 and VI.96-2 also make it clear that human vision has little or no preattentive sensitivity to the *entropy* of the probability distribution characterizing a scramble. The entropy of a probability distribution p reflects the overall "randomness" of p. Specifically, the entropy of p is the expectation of $-\log_2(p[X])$, where X is a random variable with distribution p. Thus if p assigns probability 1 to a particular value (and zero to all others; i.e., a uniform field), then p's entropy is zero. If p assigns equal probability to each of 2^n values (and zero to all others), then p's entropy is n. The entropy of the uniformly distributed scramble on the right in Figure VI.96-2 is 4.09 whereas the entropy of the three-gray-level scramble on the left is 1.48. Thus if human vision were preattentively sensitive to the entropy of a scramble's distribution, there should be a strong preattentive difference between the scrambles on the left and right in Figure VI.96-2.

To take a third example, it has been proposed that human vision contains a "boundary contour" system that registers the locations of boundaries between homogeneous regions (Grossberg & Todorović, 1988). This system is hypothesized to play a central role in the process through which brightnesses are assigned to homogeneous regions in the visual field. Suppose that human vision does indeed possess such a boundary contour system and that this system registers the existence of a boundary with a degree of strength that is invariant with respect to the contrast difference across the boundary.[1] In this case, the response of the boundary contour system to Figure VI.96-2 should resemble the image shown in Figure VI.96-4. The fact that we easily see the difference between the right and left sides of Figure VI.96-4 due to the greater number of edges on the right suggests that if human vision does embody a boundary contour system, either this system does not operate as we have hypothesized here or else the output of this system is not used to inform preattentive vision.

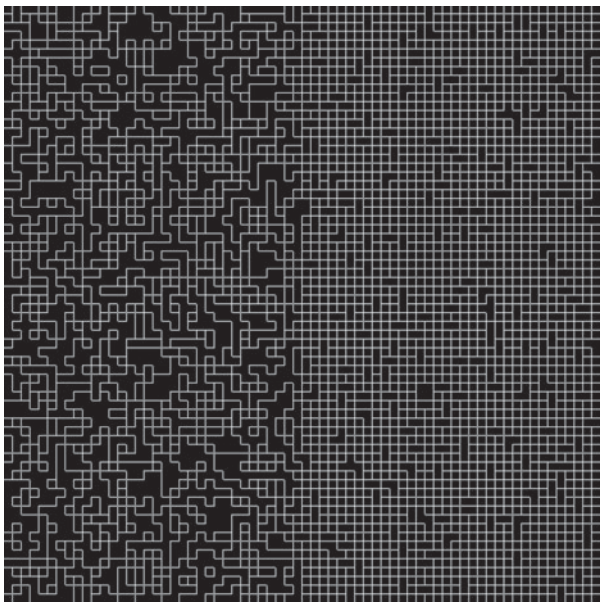


Figure VI.96-4. The boundary map of the scrambles of Figure VI.96-2. Every white line in this figure marks a change in gray level between homogeneous regions in the scrambles of Figure VI.96-2.

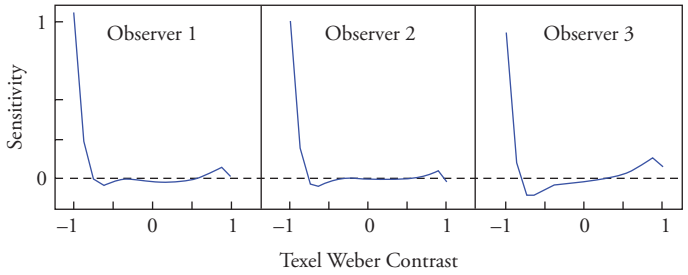


Figure VI.96-5. Sensitivity of the blackshot texture mechanism as a function of the Weber contrast of the scramble texture elements. (From Chubb et al., 2004).

HISTORY AND EXPLANATION

Figure VI.96-1 was prompted by the experiments of Chubb, Econopouly, and Landy (1994) who investigated preattentive discrimination of gray-level scrambles. In their experiments, participants judged the orientation (vertical vs. horizontal) of a briefly presented square wave whose bars were defined by scrambles with different histograms. In particular, all the scrambles used by Chubb et al. had mean and variance equal to those of the uniformly distributed scramble shown on the right in Figure VI.96-2; scramble histograms varied only in higher-order moments. They found that for scrambles in this restricted class, sensitivity to the orientation of the scramble-defined square wave could be understood in terms of a single preattentive mechanism that was differentially sensitive to very dark pixels, subsequently called the *blackshot mechanism*. Chubb et al. (1994) provided a partial characterization that sufficed to predict the differential activation of the blackshot mechanism by scrambles in the class to which their study was restricted. Chubb, Landy, and Econopouly (2004) later extended these results to complete the characterization of the blackshot mechanism.

The function that gives the sensitivity of the blackshot mechanism to different gray levels is plotted for three observers in Figure VI.96-5. As these plots make clear, the blackshot mechanism is highly activated by squares with Weber contrasts near –1 (i.e., with gray levels near black) but discriminates very poorly between texture elements with Weber contrasts greater than –0.95.

The results of Chubb et al. (1994) suggested that if two scrambles were equated in mean, variance and blackshot (although the blackshot sensitivity function was not fully characterized until 2004), then they should be preattentively indiscriminable. It was this observation that led us (Chubb, Darcy, & Sperling, 1993) to hunt down the scramble shown on the left-hand side in Figure VI.96-2. Specifically, we conducted a brute force search through the space of scrambles made of only three gray levels to see if we could find such a three-gray-level scramble with mean, variance, and blackshot equal to those of the uniformly distributed scramble.

It was not a foregone conclusion that such a scramble would exist. For any fixed set of three gray levels, once one assigns probabilities to two of the gray levels, the probability of the third gray level is fixed as well. For the same reason, once one adjusts the probabilities of a three-gray-level scramble to make the mean and variance of the scramble equal to those of the uniformly distributed scramble, the blackshot of the scramble is fixed as well. As it turned out, however, there did exist a scramble with only three gray levels whose mean, variance, and blackshot were nearly equated to those of the uniformly distributed scramble, a scramble similar to the one shown on the left side of Figure VI.96-2.

In reproducing the scramble illusion for various different demonstrations—some in printable format, others on different display devices—we have found that the illusion is easier to achieve (i.e., it is easier to adjust the probabilities of the three gray levels in the Figure VI.96-2 left-hand scramble so as to make this scramble appear identical to the uniformly distributed scramble on the right) if the second lowest gray level used in both scrambles is shifted to be closer to black. In the current display, instead of being 1/16 (which would place it midway between the first and the third gray levels), the second gray level has been set to 1/32. In the scramble on the left, all gray levels have probability zero except the second, the seventh, and the fifteenth, which have probabilities 0.1642, 0.4517, and 0.3841, respectively.

Given that we arrived at these probabilities by adjusting the proportions of gray levels numbers 2, 7, and 15 so as to make the 3- and 17-gray-level scrambles as similar in appearance as possible, it is reasonable to ask: How do the mean, variance, and blackshot of the 3- and 17-gray-level textures compare? The answer is that the 3-gray-scale scramble has 2.6% higher mean, 2.1% higher variance, and approximately 4.3% higher blackshot than the 17-gray-level scramble.

These disparities are not unexpected. Although we know that the blackshot mechanism is very sharply tuned for the darkest scramble elements, the characterization of the sensitivity function provided by Chubb et al. (2004) is

a seventh-order polynomial approximation that may not be perfectly accurate across the range of gray levels near black where the function is steep. We therefore cannot rule out the possibility that the disparity in blackshot between the 3-gray-level and 17-gray-level scrambles may be due to error in our estimate of the blackshot sensitivity function. The disparities between the means and variances of the 3- and 17-gray-level scrambles, although slight, are also consonant with previous results. As shown by Chubb, Nam, Bindman, and Sperling (2007), the preattentive sensitivity of human observers to scramble mean and variance is conferred by mechanisms that are actually sensitive to statistics that differ slightly from mean and variance.

In color theory, two patches that appear identical in color but have different wavelength compositions (different spectra) are called metamers. By analogy, two scrambles that appear preattentively identical but that have different gray-level histograms are texture metamers. Human color vision is three-dimensional. This implies that any color can be matched with an appropriate combination of three primary colors. Tridimensionality is the basis of color photography and color video. Similarly, we (Chubb et al., 2007) have found that preattentive human vision for scrambles of the type shown here is three-dimensional (excluding scrambles whose histograms have very low entropy). The analogy to color thus suggests that one should be able to match the appearance of any scramble by taking an appropriate mixture of some fixed set of three well-chosen gray levels. However, this is not quite true: as soon as one assigns the proportions of the first two gray levels, the proportion of the third gray level is also determined (because the probability distribution characterizing any scramble is constrained to sum to 1). This means that with only three gray levels, one does not have enough freedom to construct a scramble that matches a given scramble in all three of mean, variance, and blackshot. In order to ensure the ability to obtain such a match, a fixed set of not three but rather *four* gray levels is required to match the mean, variance, and blackshot of an arbitrary scramble. (Although it is possible to adjust the proportions of the three gray levels used in the left-hand scramble in Fig. VI.96-2 to match the appearance of the uniformly distributed scramble on the right-hand side of Fig. VI.96-2, there are many scrambles whose mean, variance, and blackshot cannot be matched by adjusting the proportions of these same three gray levels). In color metamers, no amount of selective attention will reveal the difference between metamers because there are only three kinds of color-sensitive receptors (retinal cones). In scramble perception, selective attention can bring into play visual processes other than the mechanisms typically used in texture perception. That is, selective attention to internal details of a texture is not a typical texture-perception process. However, for monochromatic scrambles that are perceived preattentively, texture perception is a three-dimensional process analogous to the three-dimensional perception of color patches.

We can reveal the physical difference between the scrambles in Figure VI.96-1 by applying an appropriate transformation of gray level. To see what is really going on in Figure VI.96-1, we apply a transformation that maps gray level #2 onto the value 0.83, gray level #7 onto 1.0, gray level #15 onto 0.98, and all other gray levels onto zero. This particular function has been chosen to maximize the signal-to-noise ratio in the output image between the texture element values in the region of the 3-gray-level scramble versus the region of the 17-gray-level scramble. The resulting image is shown in Figure VI.96-6. This function maps regions filled with the 3-gray-level scramble onto regions of high average value and regions filled with the 17-gray-level scramble onto regions of low average value.

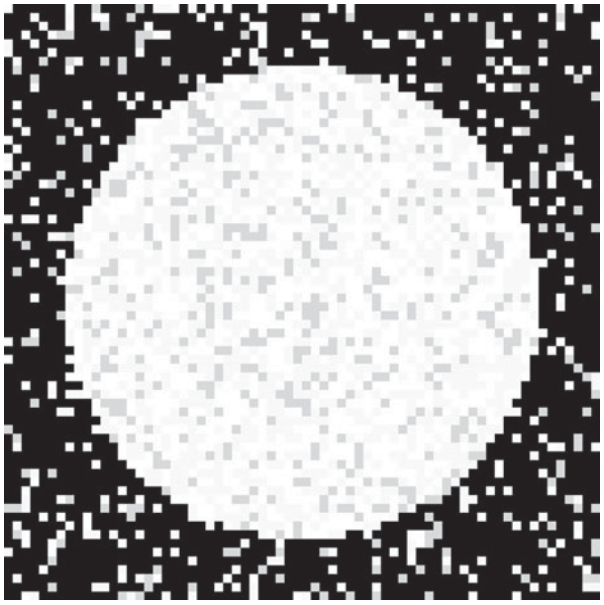


Figure VI.96-6. The organization of the 3- and 17-level scrambles in Figure VI.96-1 is revealed by applying a function to Figure VI.96-1 that maps gray level #2 onto the value 0.83, gray level #7 onto 1.0, gray level #15 onto 0.98, and all other gray levels onto zero. This function maximizes the signal-to-noise ratio in the output image between the texture element values in the region of the 3-gray-level scramble versus the region of the 17-gray-level scramble. As this transformation shows, Figure VI.96-1 is composed of a central disk of the 3-gray-level scramble set against a background of the 17-gray-level scramble.

The general point illustrated by the scramble illusion is that we cannot always see what is visible, even when we look directly at it. The physical differences between the scrambles on the left and right sides of Figure VI.96-2 are visible: When asked how many gray levels are used in the scramble on the left, we easily see that the answer is three; and when asked how many are used in the scramble on the right, we quickly affirm that the answer is substantially more. Nonetheless, we do not notice these differences without explicit prompting, which raises an interesting question: What else are we missing that is right before our eyes?

NOTE

1. In the simulations presented by Grossberg and Todorović (1988), the boundary contour system does not, in fact, have this property. Rather, boundary strength increases with the contrast difference on the two sides of the boundary.

REFERENCES

Chubb, C., Darcy, J., & Sperling, G. (1993). Metameric matches in the space of textures comprised of small squares with jointly independent intensities [Abstract]. *Investigative Ophthalmology and Visual Science*, *34*(Suppl. 4), 1289.

Chubb, C., Econopouly, J., & Landy, M. S. (1994). Histogram contrast analysis and the visual segregation of IID textures. *Journal of the Optical Society of America A*, *11*, 2350–2374.

Chubb, C., Landy, M. S., & Econopouly, J. (2004). A visual mechanism tuned to black. *Vision Research*, *44*, 3223–3232.

Chubb, C., Nam, J.-H., Bindman, D. R., & Sperling, G. (2007). The three dimensions of human visual sensitivity to first-order contrast statistics. *Vision Research*, *47*, 2237–2248.

Grossberg, S., & Todorović, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena. *Perception & Psychophysics*, *43*, 241–277.

Robson, J. G. (1980). Neural images: The physiological basis of spatial vision. In C. S. Harris (Ed.), *Visual coding and adaptability* (pp. 177–214). Hillsdale, NJ: Lawrence Erlbaum.

Chapter 97
Amodally Completed Angles

Walter Gerbino

THE GERBINO ILLUSION

When coincidentally occluded like in Figure VI.97-1, a regular hexagon appears distorted (Gerbino, 1978). Da Pos and Zambianchi (1996) labeled the effect "the Gerbino illusion" and related it to a variant of the Poggendorff illusion (Fig. VI.97-2a) studied by Burmester (1896, Fig. 15). As noticed by Lewis (1892), Pierce (1898, p. 820) and Metzger (1953, Fig. 187; 1975, Fig. 194),[1] the same distortion can be observed in a real three-dimensional setting, when a far Gothic arch is coincidentally occluded by a near pillar, as depicted in panels *b* and *c* of Figure VI.97-2. Ninio (1998, Figs. 12–14) and Seckel (2002, Fig. 91) discussed a variant of the perceptually distorted hexagon in which the occluding triangles were smaller than in Figure VI.97-1, showing that the illusion survives relevant changes in the proportion of occlusion, provided that the essential feature (i.e., the coincidence between each occluded vertex and an occluding edge) is preserved.

The educational website Cut-The-Knot (www.cut-the-knot.org/Curriculum/Geometry/Bender.shtml) supports a complex interactive video that allows observers to experience both the Gerbino illusion (involving coincidentally occluded angles) and the phenomenal rounding of symmetrically occluded angles. The dynamic demo allows users to choose several polygons as occluding and occluded shapes and makes evident that the distortion is phenomenally immediate. The distortion appears and disappears in close correspondence with the coincidental kinetic occlusion of vertices, suggesting that the main determinants of the Gerbino illusion are temporally, as well as spatially, local.

Figure VI.97-3 illustrates a variant of the Gerbino illusion in which the shape of occluding triangles has been modified to control for possible contrast/assimilation effects between intersecting segments. In Figure VI.97-3, both sides of every partially occluded angle bisect the corresponding occluding edge. The distortion is still there, showing that it cannot be attributed to a perceived mislocation of line endings induced by the asymmetrical position of T-junctions in Figure VI.97-1. Note also that changing the proportions of occlusion in *a* versus *b* produces a slight change in the perceived amount of distortion, if any. Consider the following aspects of the Gerbino illusion observed in Figures VI.97-1 and VI.97-3.

The illusion differs from the family of Poggendorff configurations involving acute/obtuse angles, where distortions may be accounted for by a regression toward right angles (Hotopf & Ollerearnshaw, 1972; Ninio & O'Regan, 1999). The Gerbino illusion is obtained in configurations in which the occluding edges intersect the partially occluded sides forming acute/obtuse angles, like in some displays by Gerbino (1978), included also in the Cut-The-Knot demo. However, the distortion is strong and clear even when all T-junctions involve only right angles, showing that it does not depend on image-level interactions between adjacent contours. Rather, it depends on amodal completion (Kanizsa & Gerbino, 1982) and, in particular, as originally proposed by Gerbino (1978), on an "irresistible" tendency of contours to continue behind occluders, consistent with the generic viewpoint assumption (Freeman, 1994). When occlusion is coincidental (like in Figs. VI.97-1, VI.97-2, and VI.97-3) even a minimal tendency of both T-junction stems to continue beyond their retinally specified positions leads to an arrangement of extended T-junction stems incompatible with a straight-line partially occluded angle. This incompatibility accounts for the global distortion of the symmetrical hexagon, as well as for the local displacement of the side owning the vertex of the angle beyond the point of rectilinear extrapolation of the other side, under the assumption that amodal continuation affects the perception of modal parts too.

The illusion is robust with respect to changes of the attentional focus. When attention is distributed over the global pattern, the partially occluded hexagon loses its regularity and looks distorted. When attention is focused on a single partially occluded angle, the two sides look locally misaligned and incompatible with the physical occlusion of a simple straight-line angle. This is consistent with the idea that the basic determinants of the effect are local: the global distortion simply results from the combination of local effects.

The local distortion occurs when an angle is partially covered by an occluder whose edge is coincident with the vertex (i.e., when occluding edge and occluded vertex are aligned along the line of sight). The effect can be observed when an isolated angle is coincidentally occluded, as well as when all angles of a regular or irregular polygon are coincidentally occluded (Gerbino, 1978). However, when

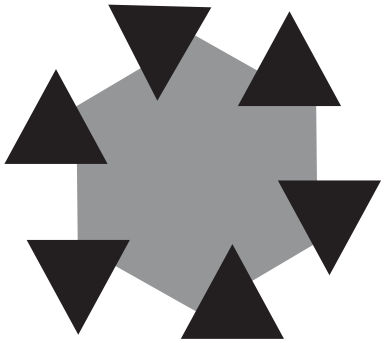


Figure VI.97-1. The distorted hexagon (Gerbino, 1978). Globally, the symmetry of the partially occluded hexagon is perturbed. Locally, each amodally completed angle looks wrong (i.e., incompatible with the continuation of both sides along a straight trajectory).

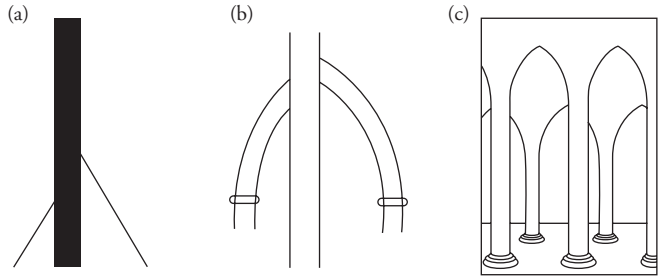


Figure VI.97-2. In the variant of the Poggendorff illusion shown in (a) (Burmester, 1896, Fig. 15), at least part of the distortion can be explained by a perceptual regression toward right angles. The displays shown in (b) (Pierce, 1898, Fig. 4) and (c) (Metzger, 1953, Fig. 187; 1975, Fig. 194) illustrate an ecological condition in which the occluding edge and the occluded vertex are aligned along the line of sight, generating a distortion of the partially occluded angle.

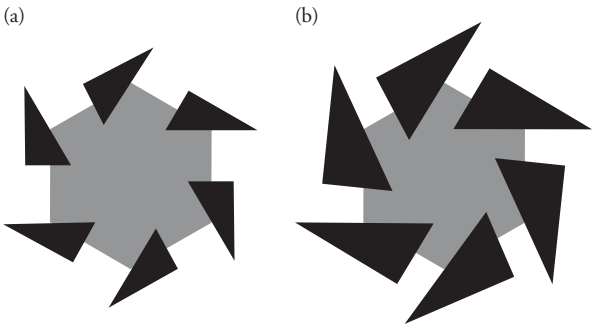


Figure VI.97-3. A variant of the Gerbino illusion in which each side of the partially occluded hexagon bisects the corresponding side of an occluding triangle, to control for possible contrast/assimilation effects between adjacent segments. The effect in (b) is stronger than in (a), showing that the distortion depends of the amount of occlusion.

the partially occluded shape undergoing the perceived distortion is a regular polygon (like the hexagon in Figs. VI.97-1 and VI.97-3), the persistence of a strong illusion becomes theoretically relevant, since it falsifies the Gestalt-grounded idea that perception always tends to global *Prägnanz* (Kanizsa & Luccio, 1986). Like in other geometric illusions, in the Gerbino illusion a proximal regularity (i.e., a good shape available in the optic input) is perceptually distorted.

AMODAL COMPLETION

Occlusion is a pervasive aspect of visual ecology. Opaque objects occlude parts of themselves and other objects, leading to a fragmentary optic input. This is the rule more often than the exception. Therefore, it is no surprise when we realize that the visual system is endowed with mechanisms supporting the perception of a world of phenomenally complete objects that go beyond input fragments. Rather, the specific shapes of phenomenally completed objects are an important source of information about the "inner forces" of perceptual organization (Koffka, 1935). Such shapes emerge in the absence of local stimulation and are likely to reflect the idiosyncratic preferences of our visual system.

To name the occluded parts of visual objects Metzger (1936, ch. 8) used the expression *unsichtbar vorhanden* (invisibly present) and discussed them as a test bench for the perceptual tendency toward a good Gestalt. Michotte and Burke (1951) introduced what they conceived as a new enigma of perceptual psychology, *le donnée amodal* (the amodal datum). They chose the term *amodal* to emphasize not the lack of local stimulation supporting the perceived shape of occluded parts but the existence of anomalous visual experiences not marked by color or other *modal* attributes of vision. The Michotte school (Michotte, Thinès, & Crabbé, 1964) used the expression *compléments amodaux* (amodal complementary parts) in a descriptive sense and clarified that both modal parts and amodal complements of surfaces are a function of stimulus conditions. Kanizsa (1955) went a step further and explicitly defined "amodal completion" as a process that integrates input fragments on the basis of self-organizing forces to produce a perceptual result that is simpler and more regular than the fragments. For a criticism of his own position, see the 1986 addendum to the English translation of the original paper on quasi-perceptual margins (Kanizsa, 1955).

In the contemporary literature, the expression "amodal completion" covers both the phenomenal presence of occluded parts and the underlying processes that lead to their formation. Within such a perspective, amodally completed angles represent a well-defined domain for testing perceptual theories and producing relevant demonstrations.

Before reviewing what we know about amodally completed angles, let us mention a related, controversial point, that is, the possibility that exactly the same processes underlie various kinds of visual completion: modal (Kanizsa's illusory figures), amodal (shapes behind occluders), and possibly virtual (constellations derived from the unification of sparse elements). Kellman and others (Albert, 2007; Kellman & Loukides, 1987; Kellman & Shipley, 1991; Kellman, Yin, & Shipley, 1998; Kellman, Garrigan, & Shipley, 2005; Kalar, Garrigan, Hilger, Wickens, & Kellman, 2010) advocated an *identity hypothesis* whose validity has been disputed on empirical and theoretical grounds (Anderson, 2007; Anderson, Singh & Fleming, 2002; Ringach & Shapley, 1996; Singh, 2004; Zhou, Tjan, Zhou, & Liu, 2008). For a discussion of perceptual completions at modal and amodal levels, see van Lier and Gerbino (2015).

At least part of the controversial evidence can be accounted for in the framework of a *weak identity hypothesis* (Massad, 2006) in which the precise shape of completed angles depends on the specific value of a free parameter, which in the field model by Fantoni and Gerbino (2003) is

the *relative strength* of two completion factors: *good continuation* and *minimal path*. This idea introduces us to angle interpolation, that is, to the discussion of underlying processes at the contour level, leading to the completion of contour fragments. Such processes, possibly modulated by other completion processes active at the level of surfaces and volumes, seem crucial for the explanation of the Gerbino illusion.

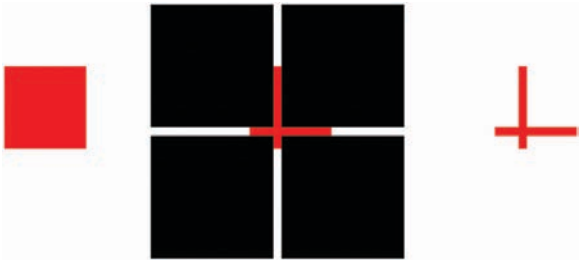
ANGLE INTERPOLATION

Let us shift our attention from the Gerbino illusion to the generic case of noncoincidentally occluded angles—more precisely, of angles whose straight sides are partially occluded by edges intersected at right angles—to avoid the confounding effect of local contour interactions. Even when completion of both sides by straight-line extrapolation would lead to the perception of a regular polygon bounded by straight contours, the amodal trajectory is perceived as rounded. The phenomenal rounding has been observed in a variety of conditions and fits a generic perceptual preference toward a smooth completion over a solution involving a discontinuity (i.e., an abrupt change of direction).

Hempstead (1901, p. 188) reported that shapes with sharp angles appear rounded at low contrast. More interestingly, Rosenbach (1902; see also Metzger, 1936, Fig. 141) observed rounding in what are now called the moving visual phantoms (Tynan & Sekuler, 1975). The Rosenbach effect is a case of apparent transparency (i.e., of modal completion) observed in conditions of physical occlusion that normally lead to amodal completion. The rounding observed by Rosenbach captures a basic feature of angle interpolation in both modal and amodal completion (Fantoni & Gerbino 2003; Gerbino 1978; Singh, 2004; Zhou et al., 2008). In modal completions, the illusory blob perceived in the Koffka cross—so called despite the fact that Koffka (1935) introduced it for a different purpose—appears rounded (Landauer, 1978), and only the addition of anchoring dots supports the perception of a square with straight sides and sharp angles (Day, 1987; Day & Jory, 1980). Most observers perceive amodally completed angles as rounded even when they belong to a simple and regular form like a square. Figure VI.97-4 illustrates two positions of a square seen through vertical and horizontal apertures between occluders. The amodally completed blobs perceived in *a* and *b* look different, since the interpolated shape of each corner depends on the local extrapolation of visible sides.



Figure VI.97-4. Two frames of a partially occluded (physical) square moving behind occluders while keeping its orientation constant and never revealing its vertices. Each occluded angle appears differently rounded, depending on local conditions. The dominance of local determinants of interpolation leads to different perceived shapes of the partially occluded object.



Video VI.97-1. The central pattern is perceived as a nonrigid blob, despite being originated by the occlusion of a revolving square (identical to the one on the left). The red patterns on the center and on the right are geometrically identical, but are perceived in quite different ways: the one on the center as a nonrigid blob moving counterclockwise; the one on the right as two bars sliding clockwise one over the other.

The "eccentric square" shown in the center of Video VI.97-1 follows the logic illustrated in Figure VI.97-4. The comparison with the pattern on the right is based on well-known motion effects (Lorenceau & Shiffrar, 1992; McDermott, Weiss, & Adelson, 2001). On the right one sees two red bars, sliding one over the other in relative counterclockwise motion, like in the chopstick illusion (Anstis, 1990). In the center one may see two bars moving exactly like on the right, but the pattern is easily reorganized and perceived as a nonrigid blob moving clockwise. The perceptual shift is explained by grouping of two short vertical segments oscillating horizontally and two short horizontal segments oscillating vertically into a common-motion structure, against the figural unity of the vertical and horizontal bars visible on the right. However, from the point of view of amodal completion, the important aspect is the nonrigidity of the blob, consistent with the local interpolation of partially occluded angles, which generates slightly different curvilinear trajectories behind each occluding square. Also for the eccentric square one might expect the dominance of a tendency toward maximal regularity, that in this case would support the perception of a rigid square (maybe rounded) in constant motion. On the contrary, like in the static case of the hexagon of Figures VI.97-1 and VI.97-3, local factors dominate, leading to the loss of rigidity.

The trajectory of completed angles (i.e., the amount of phenomenal rounding) has been measured using two methods: (a) the probe localization technique (Fantoni, Bertamini, & Gerbino, 2005; Gerbino & Fantoni, 2006; Guttman & Kellman, 2004; Kellman, 2000), in which a dot or short segment is flashed in the region crossed by the perceptually interpolated trajectory and observers tell whether the probe is inside/outside the completed shape; and (b) matching with a fully visible comparison shape (Fantoni et al., 2005; Singh, 2004). Singh argued that the probe localization technique may introduce undesirable artifacts in the measurement of interpolated contour trajectories. Such artifacts, however, do not undermine the basic rounding effect and seem relevant only for the estimation of the exact amount of phenomenal rounding.

Various models predict the precise shape of the amodal trajectory of partially occluded angles (for reviews see Ben-Yosef & Ben-Shahar, 2010; Fantoni & Gerbino, 2003; Fulvio, Singh, & Maloney, 2015; Massad, 2006). In general, they generate smooth curvilinear solutions that can be conceived as the output of an interpolation mechanism activated by the presence of incomplete fragments and driven

toward closure of the trajectory. In many cases the difference between the predictions of various models is small compared to the empirical difficulty of estimating amodal trajectories. Therefore, the following discussion refers only to main aspects of the field model by Fantoni and Gerbino, which provides a useful framework for explaining both the rounding of amodal angles under generic occlusion and the Gerbino illusion in the limiting case of coincidental occlusion.

According to the field model (Fantoni et al., 2005; Fantoni & Gerbino, 2003; Gerbino & Fantoni, 2006), the smooth curvilinear trajectory observed in generic cases of angle occlusion is generated by the chaining of vector sums derived from the combination of two fields induced by input fragments: the minimal path field and the good continuation field. The model implements the interaction between neural units sensitive to orientation, supporting the good continuation of contours, and units sensitive to endpoint proximity, supporting closure along the minimal path. Through the change of a single free parameter, the relative strength of good continuation and minimal path, the model embodies the effects of a number of variables known to affect the interpolation trajectory, namely

- *scale dependence* (Gerbino & Fantoni, 2006), leading to flatter trajectories as the retinal gap between endpoints decreases;
- *support ratio* (Shipley & Kellman, 1992), leading to flatter trajectories as the ratio of physically specified length to total side length decreases;
- *orientation* (Fantoni & Gerbino, 2003), leading to flatter trajectories as the direction of the minimal path becomes closer to cardinal axes, since the strengths of both good continuation and minimal path increase as their direction gets closer to cardinal axes;
- *contour curvature polarity* (Fantoni et al., 2005), leading to flatter trajectories when the occluded shape is locally convex, rather than concave; and
- *global shape* (Fantoni & Gerbino, 2003), supporting amodal trajectories that maximize the global symmetry of the occluded shape.

The first three variables refer to interpolation at the contour level, while the last two refer to interpolation at the surface level. The dependence of the precise shape of amodal completion on retinal size, measured by Gerbino and Fantoni (2006), may appear surprising since, in general, scale invariance is a desirable property of perception. Size illusions connected to size constancy suggest that perceived size tends to be invariant with respect to changes of retinal size. However, one can easily demonstrate that, for instance, the partially occluded blob in Figure VI.97-4a looks more similar to a square when observed from a short distance (large retinal gap) and more similar to a disk from a far distance (small retinal gap). This implies that proximal constraints play a significant role in the process of amodal completion, which in this respect is more similar to retinotopic filling-in processes (Pessoa, Thompson, & Noë, 1998) than to a type of categorical perception supporting the recognition of partially occluded objects despite missing data. The scale dependence of amodally completed angles parallels the scale dependence of modal illusory contours, easily demonstrated in patterns like the Koffka cross. For a discussion and a simple distortion effect see Gerbino and Kanizsa (1987).

INTERPOLATION VERSUS APPROXIMATION

With the notable exception of Fantoni, Gerbino, and Kellman (2008) and Fulvio, Singh, and Maloney (2008), the amodal completion of partially occluded angles has been conceived as the production of a trajectory that interpolates two veridically represented input segments. This is the case also for the original explanation of the Gerbino illusion, which attributes the distortion to the inclusion into the perceived shape of both the veridical modal parts and the amodal additions—influenced by good continuation—that disrupt the regularity of the occluded shape (Gerbino, 1978).

As suggested by Fantoni et al. (2008), a radically different explanation might be proposed following the idea that amodal completion does not involve visual interpolation but visual approximation. The difference between such notions becomes clear as soon as one considers the problem of fitting a polynomial function to a set of points (Ullman, 1996, pp. 141–143): interpolation generates a curve that connects all points, minimizing the changes of direction, while approximation generates a curve that minimizes the distances from the points, without exactly touching them all. Approximation involves an error (increasing as the degree of the fitting function decreases) but can effectively account for noisy data.

In the case of angle completion: (a) interpolation implies the literal representations of the two input fragments and the generation of a smooth nonmonotonic contour connecting them, while (b) approximation implies the generation of a smooth hexagon that cannot perfectly match the arrangement of input fragments—given the coincidental occlusion of vertices—and the emergence of the feeling of the mismatch between input evidence and the internally generated shape. For evidence supporting approximation in amodal completion, see Fantoni et al. (2008).

As a possible explanation of the Gerbino illusion, approximation accounts (better than interpolation) for the perceived mislocation of *modal* sides of each partially occluded angle, under the assumption that perception represents not only the optic input but also the degree of fitness of internal object models to such an input.

NOTE

1. The Gothic arc illusion, not included in the first edition of Metzger's *Gesetze des Sehens* (1936), is discussed in the second (1953) and third (1975) expanded editions. English translation of the first edition: *The Laws of Seeing* (Cambridge, MA: MIT Press, 2006).

REFERENCES

Albert, M. (2007). Mechanisms of modal and amodal interpolation. *Psychological Review*, *114*, 455–469.

Anderson, B. L. (2007). The demise of the identity hypothesis and the insufficiency and nonnecessity of contour relatability in predicting object interpolation: Comment on Kellman, Garrigan, and Shipley (2005). *Psychological Review*, *114*, 470–487.

Anderson, B. L., Singh, M., & Fleming, R. W. (2002). The interpolation of object and surface structure. *Cognitive Psychology*, *44*, 148–190.

Anstis, S. (1990). Imperceptible intersections: the chopstick illusion. In A. Blake & T. Troscianko (Eds.), *AI and the eye* (pp. 105–117). New York: Wiley.

Ben-Yosef, G., & Ben-Shahar, O. (2010). A biologically inspired theory for non-axiomatic parametric curve completion. In *Proceedings of the Asian Conference on Computer Vision* (pp. 1–14). New York: Springer.

Burmester, E. (1896). Beitrag zur experimentellen Bestimmung geometrisch-optischer Täuschungen. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, *12*, 355–394.

Da Pos, O., & Zambianchi, E. (1996). *Illusioni ed effetti visivi. Una raccolta (Visual illusions and effects. A collection).* Milan: Guerini e Associati. http://www.guerini.it/index.php/illusioni-ed-effetti-visivi-una-raccolta-visual-illusions-and-effects-a-collection.html

Day, R. H. (1987). Cues for edge and the origin of illusory contours: An alternative approach. In S. Petry & G. Meyer (Eds.), *The perception of illusory contours* (pp. 246–252). New York: Springer.

Day, R. H., & Jory, M. K. (1980). A note on a second stage in the formation of illusory contours. *Perception & Psychophysics*, *17*, 89–91.

Fantoni, C., Bertamini, M., & Gerbino, W. (2005). Contour curvature polarity and surface interpolation. *Vision Research*, *45*, 1047–1062.

Fantoni, C., & Gerbino, W. (2003). Contour interpolation by vector field combination. *Journal of Vision*, *3*, 281–303. http://jov.arvojournals.org/article.aspx?articleid=2121569

Fantoni, C., Gerbino, W., & Kellman, P. J. (2008). Approximation, torsion, and amodally completed surfaces. *Vision Research*, *48*, 1196–1216.

Freeman, W. T. (1994). The generic viewpoint assumption in a framework for visual perception. *Nature*, *368*, 542–545.

Fulvio, J., Singh, M., & Maloney, L. (2008). Precision and consistency of contour interpolation. *Vision Research*, *48*, 831–849.

Fulvio, J., Singh, M., & Maloney, L. (in press). Visual interpolation and extrapolation of contours. In S. Gepshtein, L. Maloney, & M. Singh (Eds.), *Oxford handbook of computational perceptual organization*. Oxford: Oxford University Press. http://www.psych.nyu.edu/maloney/FulvioSinghMaloney.2015.pdf

Gerbino, W. (1978). Some observations on the formation of angles in amodal completion. *Italian Journal of Psychology*, *5*, 88–100.

Gerbino, W., & Fantoni, C. (2006). Visual interpolation is not scale invariant. *Vision Research*, *46*, 3142–3159.

Gerbino, W., & Kanizsa, G. (1987). Can we see constructs? In S. Petry & G. Meyer (Eds.), *The perception of illusory contours* (pp. 246–252). New York: Springer.

Guttman, S. E., & Kellman, P. J. (2004). Contour interpolation revealed by a dot localization paradigm. *Vision Research*, *44*, 1799–1815.

Hempstead, L. (1901). The perception of visual form. *The American Journal of Psychology*, *12*, 185–192.

Hotopf, W. H. N., & Ollerearnshaw, C. (1972). The regression to right angles tendency and the Poggendorff illusion: II. *British Journal of Psychology*, *63*, 369–379.

Kalar, D., Garrigan, P., Hilger, J., Wickens, T., & Kellman, P. J. (2010). A unified model for contour interpolation. *Vision Research*, *50*, 284–299.

Kanizsa, G. (1955). Margini quasi-percettivi in campi con stimolazione omogenea. *Rivista di Psicologia, 49*, 7–30. (English translation: Quasi-perceptual margins in homogeneously stimulated fields [with a 1986 addendum], in S. Petry & G. E. Meyer [Eds.], *The perception of illusory contours* [pp. 40–49]. New York: Springer, 1987)

Kanizsa, G. (1979). *Organization in vision*. New York: Praeger.

Kanizsa, G., & Gerbino, W. (1982). Amodal completion: Seeing or thinking? In J. Beck (Ed.), *Organization and representation in perception* (pp. 167–190). Hillsdale, NJ: Lawrence Erlbaum.

Kanizsa, G., & Luccio, R. (1986). Die Doppeldeutigkeiten der Prägnanz. *Gestalt Theory*, *8*, 98–135. (Italian translation: Le ambiguità della pregnanza, in G. Kanizsa [Ed.], *Vedere e pensare* [Chapter IV]. Bologna: Il Mulino, 1991)

Kellman, P. J. (2000). An update on Gestalt psychology. In B. Landau, J. Sabini, J. Jonides, & E. L. Newport (Eds.). *Perception, cognition, and language: Essays in honor of Henry and Lila Gleitman* (pp. 157–190). Cambridge, MA: MIT Press.

Kellman, P. J., Garrigan, P., & Shipley, T. F. (2005). Object interpolation in three dimensions. *Psychological Review*, *112*, 586–609.

Kellman, P. J., & Loukides, M. G. (1987). An object perception approach to static and kinetic subjective contours. In S. Petry & G. E. Meyer (Eds.), *The perception of illusory contours* (pp. 151–164). New York: Springer.

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, *23*, 141–221.

Kellman, P. J., Yin, C., & Shipley, T. F. (1998). A common mechanism for illusory and occluded object completion. *Journal of Experimental Psychology*, *24*, 859–869.

Koffka, K. (1935). *Principles of Gestalt psychology.* New York: Harcourt, Brace.

Landauer, A. A. (1978). Subjective states and the perception of subjective contours. In J. P. Sutcliffe (Ed.), *Conceptual analysis and method in psychology: Essays in honour of W. M. O'Neil* (pp. 142–146). Sydney, Australia: Sydney University Press.

Lewis, R. T. (1892). Optical illusions (letters to editors). *Nature*, *47*, 31.

Lorenceau, J., & Shiffrar, M. (1992). The influence of terminators on motion integration across space. *Vision Research*, *32*, 263–273.

Massad, A. (2006). A perceptual grouping approach for visual interpolation between good continuation and minimal path using tensor voting. In M. Valstar, A. French, & T. Pridmore (Eds.), *Proceedings of the British Machine Vision Conference* (pp. 639–648). Durham, UK: BMVA Press.

McDermott, J., Weiss, Y., & Adelson E. H. (2001). Beyond junctions: Nonlocal form constraints on motion interpretation. *Perception*, *30*, 905–923.

Metzger, W. (1936). *Gesetze des Sehens*. Frankfurt: Kramer.

Michotte, A., & Burke, L. (1951). Une nouvelle énigme de la psychologie de la perception: le "donnée amodal" dans l'experience sensorielle. In *Actes du XIII Congrés Internationale de Psychologie* (pp. 179–180). Rome: V. Ferri.

Michotte, A., Thinés, G., & Crabbè, G. (1964). *Les compléments amodaux des structures perceptives*. Louvain, Belgium: Publications Universitaires.

Ninio, J. (1998). *La science des illusions*. Paris: O. Jacob. (English translation: *The science of illusions*, Ithaca, NY: Cornell University Press, 2001)

Ninio, J., & O'Regan, J. K. (1999). Characterisation of the misalignment and misangulation components in the Poggendorff and corner-Poggendorff illusions. *Perception, 28*, 949–964.

Pessoa, L., Thompson, E., & Noë, A. (1998). Finding out about filling-in: A guide to perceptual completion for visual science and the philosophy of perception. *Behavioral & Brain Sciences*, *21*, 723–802.

Pierce, A. H. (1898). Geometrical optical illusions. *Science*, *8*, 814–829.

Ringach, D. L., & Shapley, R. (1996). Spatial and temporal properties of illusory contours and amodal boundary completion. *Vision Research*, *36*, 3037–3050.

Rosenbach, O. (1902). Zur Lehre von den Urtheilstäuschungen. *Zeitschrift für Psychologie*, *29*, 434–448.

Seckel, A. (2002). *More optical illusions*. London: Carlton Books.
Shipley, T. F., & Kellman, P. J. (1992). Strength of visual interpolation depends on the ratio of physically specified to total edge length. *Perception & Psychophysics*, *52*, 97–106.
Singh, M. (2004). Modal and amodal completion generate different shapes. *Psychological Science*, *15*, 454–459.
Tynan, P., & Sekuler, R. (1975). Moving visual phantoms: A new contour completion effect. *Science, 188*, 951–952.
Ullman, S. (1996). *High-level vision: Object recognition and visual cognition*. Cambridge, MA: MIT Press.
van Lier, R., & Gerbino, W. (2015). Perceptual completions. In J. Wagemans (Ed.), *Oxford handbook of perceptual organization* (pp. 294–320). Oxford: Oxford University Press.
Zhou, J., Tjan, B. S., Zhou, Y., & Liu, Z. (2008). Better discrimination for illusory than for occluded perceptual completions. *Journal of Vision*, *8*, 1–17.

Chapter 98
Subjective Contours

Barbara Gillam

INTRODUCTION

Contours perceived as physically present ("modal") in a location where there are no variations in luminance or color are known as "subjective" or "illusory" contours. Figure VI.98-1 presents some examples. Because they are somewhat mysterious, subjective contours are a source of fascination to advertisers and logo designers and of course vision scientists. Every variety of visual theory tends to consider them the *piece de resistance* of their particular approach. However, there is general agreement among virtually all models that subjective contours are related to perceived occlusion.

It is important to note that subjective contours can only be considered "illusory" on the view that we see the retinal image. Once it is understood that vision is adapted to perceiving the arrangement and properties of surfaces in a scene, subjective contours can be understood as part of a process of surface recovery based on contextual information. The presence of subjective contours shows that contour delineation can be based on contextual information when luminance contours are absent. Importantly, this information not only indicates the presence of a contour where no luminance gradient is present but also the surface to which it belongs. Subjective contours are always seen as the boundaries of a nearer surface. Subjective contours thus are more informative than luminance contours alone. It seems likely therefore that the contextual process that gives rise to subjective contours often accompanies luminance contours, providing information about surface depth order. I propose that subjective contours are not *generated* when luminance information is removed but *revealed* when this occurs.

It is commonly said that subjective contours are there to "fill in" gaps that occur naturally in the images of contours when shadows or lighting conditions reduce local luminance contrast. However, subjective contours do not automatically fill in gaps, even in collinear contours. Under what conditions then do they do so? I begin to answer this question by considering the most famous example, Figure VI.98-1b, the Kanizsa figure (1955, 1979), which generated an explosion of interest in a long-dormant phenomenon (Schumann 1904; von Szily 1921).

The subjective contours in Figure VI.98-1b are seen as completing the boundary of a triangle in front of and partly hiding (occluding) three circles. Perceiving an occluding triangle "accounts for" the missing parts of the circles. *It does not fill them in to make complete circles*. Scrutiny of many examples reveals that subjective contours only fill in gaps as part of an apparently occluding surface. To understand subjective contours, then, it is necessary to understand the perception of occlusion, which is defined as seeing one surface in front of another. This involves consideration of three-dimensional (3D) spatial layout as Coren (1972) first pointed out. Information about surface occlusion and 3D layout is present in two-dimensional pictures (pictorial vision) in 3D binocular vision and in motion perception. However, the information supporting occlusion and thus subjective contours is different in each case. In this chapter I consider pictorial and binocular vision.

SUBJECTIVE CONTOURS IN PICTORIAL VISION

The best-known examples of subjective contours occur in drawings and pictures. The essential occlusion issue in pictorial vision is *border ownership*. Because overlapping surfaces in the scene are adjoined in the picture plane, the visual system has to find some basis for determining which surface is in front so that the common border can be correctly assigned to that surface and removed from the perceived shape of the further surface (for which it is irrelevant).

Figure VI.98-2a shows a variant of the Kanizsa figure with a subjective square. Figure VI.98-2b shows the same figure with the square fully delineated. In both cases the square appears to occlude the four circles. We can notice several things about the fully delineated square (Fig. VI.98-2b) relevant to the perception of occlusion. First, the occluding square forms T-junctions where it intersects the circles. By a T-junction I mean two contours that form a T where they intersect. T-junctions are generally regarded as strong cues to the direction of occlusion with the tops of the Ts typically being the *occluding* contours and the stems the *occluded* contours. In Figure VI.98-2b there are eight T-junctions, with the square forming the tops of the Ts, which cooperate to form an occluding surface. The interrupted outlines of the circles are the stems of Ts and appear occluded. The assignment of the notches to a nearer surface (the square) frees the partial circles from perceptually terminating at the notches

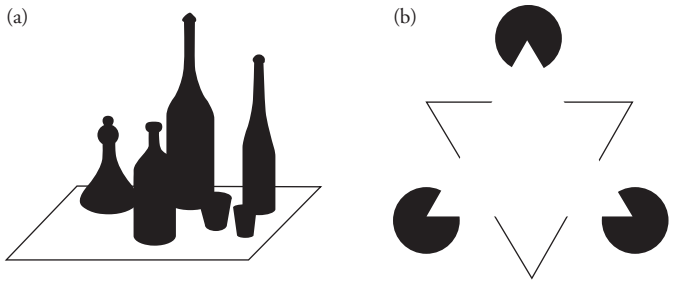


Figure VI.98-1. Examples of subjective contours by Kanizsa (1955). (a) Subjective contours are seen along the edges of the nearest bottle and the cups. (b) The classic Kanizsa figure in which a white rectangle is largely delineated by subjective contours.

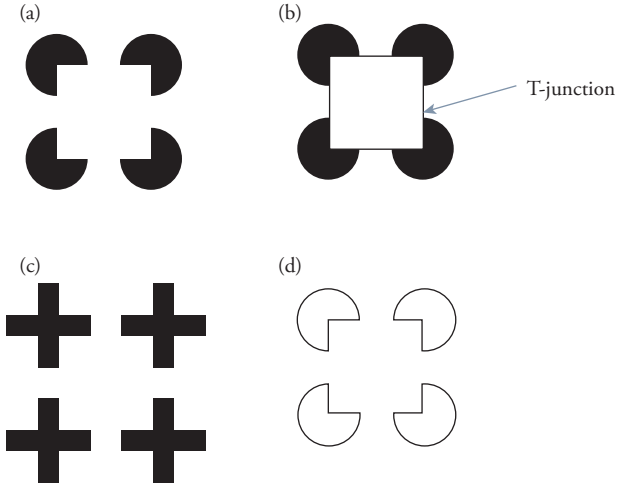


Figure VI.98-2. Variants on a Kanizsa-type figure (a). Versions of (c) and (d) can be found in Kanizsa (1979).

and allows them to perceptually continue behind the square but "amodally" (i.e., without apparent physical presence). Now look at the subjective occluding square in Figure VI.98-2a. It behaves just like the delineated occluding square in Figure VI.98-2b. It takes ownership of the edges of the notches and completes them (with the addition of subjective contours) into an occluding surface. Yet there are no T-junctions in this case to signal occlusion.

So what is the basis for the partial circles giving up border ownership at the notches? There are two factors that seem to be responsible. One is that the notches have properties that make them candidates for being *extrinsic to* (i.e., not belonging to) the circles. One might argue that the edges of the notches are potentially extrinsic because they are straight while the rest of the shape's contour is curved, that they form concavities in the convex shapes, that the notches are at odds with the completeness of the otherwise regular circles, or all of the above. A second and related factor (as stated earlier) is that several potentially extrinsic contours are aligned with each other. The potentially extrinsic edges of the notches are only seen as fully extrinsic when they can link up with other potentially extrinsic edges consistent with a common occluding surface. This is then completed with subjective contours as part of a near surface. Kellman and Shipley (1991) have formalized some of these ideas. In their model the essential characteristics for seeing contour interpolation, such as occurs in forming subjective contours, are (a) the presence of a sharp discontinuity in the direction taken by a contour (in Fig. VI.98-1b this occurs where the outlines of the circles meet the notches) and (b) the "relatability" of several such discontinuous contours. Two contours are said to be relatable if their extensions would intersect at an angle greater than 90°.

These conditions seem to describe the formation of the subjective contours in Figure VI.98-2a in which the notch contours form sharp discontinuities with the circles and are multiply relatable to each other. But are these conditions sufficient? Figure VI.98-2c has crosses with the same notches consistent with the same occluding square. It satisfies Kellman and Shipley's (1991) conditions for interpolation but has very weak subjective contours. Why? Kanizsa (1979) argued that the crosses are "self-sufficient." Another way of putting it is that the edges of the notches in this case are strongly bound to the other contours of the crosses by parallel or orthogonal relationships and have little tendency to be seen as extrinsic to them.

So it is necessary in explaining subjective contours to add to the principle that there must be sharp discontinuities in the directions of surface edges and these edges must be relatable. There must also be weak grouping of these edges with the rest of their shapes so that they are susceptible to losing ownership to an occluding surface along with other (relatable) edges.

Another striking example of the failure to see a subjective contour is shown in Figure VI.98-2d. Here the notched circles are outlined. Subjective contours extrapolate between the edges of solid shapes but never between the edges of outlined shapes. These are always seen as intrinsic to (belonging to) those shapes. Although outlines do not occur around shapes in the real world, apparently their presence triggers the code the visual system uses to indicate a contour bounding and defining a shape.

The principles I have outlined so far can be demonstrated using subjective contour inducers very different from the Kanizsa figure. These also provide new insights into the genesis of subjective contours and their relationship to occlusion. Figure VI.98-3c (based on Gillam & Chan, 2002) shows a powerful subjective contour along the aligned bottom edges in a set of five trapezia.

The subjective contour bounds an apparently occluding surface whose lower side is indeterminate (not an uncommon occurrence in the environment). On the other hand, Figure VI.98-3a shows five rectangles also aligned on their bottom edges with the same contact with the alignment as Figure VI.98-3c—yet now there is no subjective contour and little sense of occlusion at the alignment. Both figures satisfy Kellman and Shipley's (1991) principles. The bottom edges have sharp discontinuities and are relatable. There are two differences between Figure VI.98-3a and VI.98-3c that seem to be critical. One is the extrinsic edge factor that applied in the Kanizsa figure. The bottom edges in Figure VI.98-3c do not conform to the otherwise rectilinear shapes. In the case of the rectangles in Figure VI.98-3a, on the other hand, the bottom edges are bound to their shapes by parallelism and orthogonality, making it less likely that they will be seen as extrinsic and link up into an

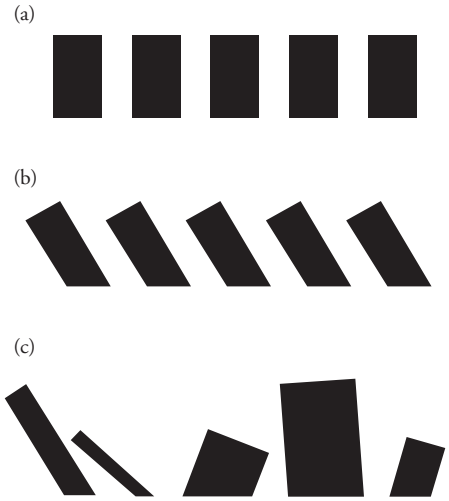


Figure VI.98-3. A figure adapted from Gillam and Chan (2002) showing three sets of five shapes aligned along the bottom. All have the same total width at the bottom alignment supported by the same amount of physical contour. The subjective contour is very strong in (c) and nonexistent in (a). In (b) it is of intermediate strength.

occluding surface. However Figure VI.98-3b also has potentially extrinsic edges. It clearly does have a stronger subjective contour than Figure VI.98-3a, but it is not nearly as strong as the one in Figure VI.98-3c. There must be another factor at work. This factor, not encountered before, is the entropy or disorder of the shapes. The trapezia in Figure VI.98-3c differ in orientation, length, thickness, and spacing. The regular alignment of these shapes along one edge is in sharp contrast to their general disorder. Such "entropy contrast" could only occur as the result of occlusion (an occluding surface is indifferent to what it occludes) and is clearly a strong indicator that occlusion is present. Two factors thus cooperate to produce the very strong subjective contour at the alignment in this figure.

Yet another genre of subjective contours beginning with the Ehrenstein figure (Fig. VI.98-4a) are those seen along thin line ends. Again the subjective contours can best be understood by considering the occlusion implications of the inducing stimuli.

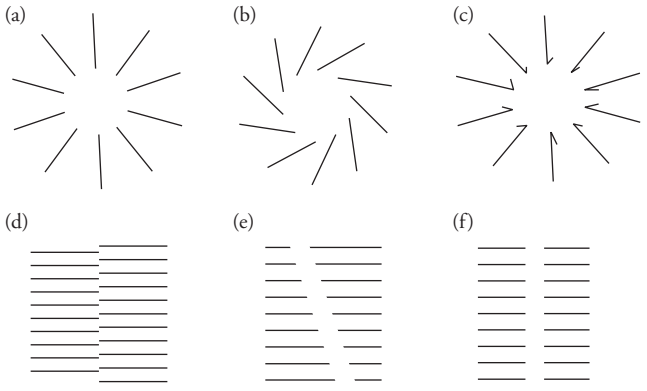


Figure VI.98-4. Examples of thin line inducers of subjective contours. (a) is an example of an Ehrenstein (1941) figure. (d) is an abutting line figure. The subjective contours in (a) are eliminated when the inducers change orientation as in (b) or when ticks are added at the endpoints as in (c).

Seeing an occluding surface with subjective contours at the line ends will occur only if the line *terminations* are seen as extrinsic to the lines, thus potentially the edge of an occluding surface. A single line end cannot appear extrinsic. It is the arrangement of line ends relative to properties of the lines that causes them to appear extrinsic. As with subjective contours based on the alignment of the edges of solid shapes, this occurs when there is weak grouping of the line ends and the lines. In the Ehrenstein illusion (Fig. VI.98- 4a), the lines are straight and radial whereas the line ends form a circle, seen as occluding the lines. However, as Kennedy (1978) showed (as in Fig. VI.98-4b), the subjective contour is much diminished when each line has an oblique orientation relative to the outline of the circle. Kennedy attributes this to the greater brightness enhancement of the circle when surrounded by the radial lines, a topic mentioned later. Another factor that strongly reduces the magnitude of subjective contours in the Ehrenstein configuration is the addition of arrowheads to the ends of the lines (Albert & Hoffman, 2000). In this case (see Fig. VI.98-4c), the perception of a subjective surface cutting off the lines is entirely eliminated. Ticks or blobs at the ends of the lines tag the lines as ending in a shape change rather than an occlusion.

The finding that Figure VI.98-4b has weaker subjective contours than Figure VI.98-4a has been taken as support for the view that subjective contours are strongest when end alignment is orthogonal to the lines, a claim that is basic to a number of physiological models. However, it is not a general principle that orthogonal contours form the strongest subjective contours. For example, the subjective contours are stronger in Figure VI.98-4e where the alignment is *not* orthogonal to the lines than in Figure VI.98-4f where they are orthogonal. A nonorthogonal alignment is more likely to be due to occlusion just because it does not have a special relationship to the lines. Another well-known example of subjective contours at thin line ends used in physiological research on subjective contours beginning with von der Heydt and Peterhans (1989) occurs when two sets of line ends are abutting (Fig. VI.98-4d) In this case the direction of occlusion is ambiguous but occlusion is supported by the common locus of the terminations.

Under certain conditions two relatable bars on either side of a larger surface can link to form an apparent transparent occluding surface with subjective contours across the larger surface. Also, as mentioned in connection with the Ehrenstein illusion (Fig. VI.98-4a), a surface defined by subjective contours can appear brighter than a background of the same luminance. However, brightness contrast does not really provide an explanation for seeing subjective contours. In Kanizsa-type figures, the brightness enhancement does not depend on the amount of contrasting luminance in the surround, and versions of the Ehrenstein figure survive the elimination of brightness contrast. Furthermore, subjective contours along a single alignment (Fig. VI.98-3c) or abutting lines (Fig. VI.98-6) occur strongly without an accompanying brightness effect. The basis of the brightness enhancement is still not understood.

BINOCULAR VISION AND SUBJECTIVE CONTOURS

Binocular organization supersedes monocular organization in producing subjective contours. In stereoscopic vision it is not the properties of the picture plane or the retinal image that determine spatial relations but the 3D layout given by the stereo view. Insofar as subjective contours are a product of the recovery of surface layout, it is not surprising that they are influenced by the introduction of stereo data.

There are three distinct aspects of stereo influence on subjective contours.

1. Stereo can bring regions of a stimulus into near depth relative to neighboring regions of the same luminance. When relatability is present for several such near regions, very strong subjective contours emerge. In Figure VI.98-5 there are hidden triangles at different separations (disparities) in the two eyes' views. There are no subjective contours in either monocular view. However, binocularly the edges of the two triangles appear to separate in depth, and strong subjective contours are seen where the contours are monocularly camouflaged.
2. Stereo can be selectively added to monocular arrangements to test hypotheses concerning the origin of subjective contours in these arrangements. Take, for example, the well-known abutting thin line figures (Fig. VI.98-4d). A stereo version is shown in Figure VI.98-6. (Gillam & Nakayama, 2002).

   In each monocular image a subjective contour is seen along the abutment. Stereoscopically, however, the upper sticks in each set of lines (we call these the forest) vary in depth from one another, whereas the lower sticks (we call these the plane) are all at the same depth. Cross-fusing the upper pair of images places the plane nearer than the forest. The subjective contour now becomes very strong. Cross-fusing the lower pair places the forest nearer than the plane. The subjective contour that was visible monocularly *now disappears*. We attribute these findings to the fact that the plane is consistent with a surface and thus is a potential occluder, while the forest is not. The forest alone does not give rise to a subjective contour even though its aligned edge provides strong entropy contrast. There also needs to be support for the presence of an occluder. Stereo allows for even more illuminating manipulations. In monocular abutting line figures such as Figure VI.98-4d, interdigitating the two sets of lines kills the subjective contour. However, interdigitating the forest/plane stereo figure (with the plane nearer) does not. (Gillam & Nakayama, 2002). The subjective contour retains its strength at small overlaps, always appearing at the edge of the forest; the plane appears to end below the ends of its sticks in order to occlude the forest at its edge. These findings provide very powerful evidence that subjective contours must be modeled at the level of determining surface layout and not at the level of encoding monocular image properties.
3. Most binocularly viewed scenes contain monocular regions. When one surface is nearer than and overlapping a more distant surface in the visual field, regions of the background surface to the left of the occluder are not visible to the right eye and vice versa. This geometry was already known to Leonardo da Vinci. Nakayama and Shimojo (1990) made a startling discovery. When monocular elements are present in a binocular field but with no occluding surface to account for them, a phantom occluding surface appears that "accounts for" the presence of the monocular elements. Anderson (1994) has shown that such subjective contours are particularly powerful when an oblique occlusion is simulated binocularly for a set of vertical lines. This is characterized by vertical differences in the position of the cut-off of the vertical lines in the two eyes. The stereograms in Figure VI.98-7 show that these vertical differences are sufficient to produce a strong subjective contour in depth.

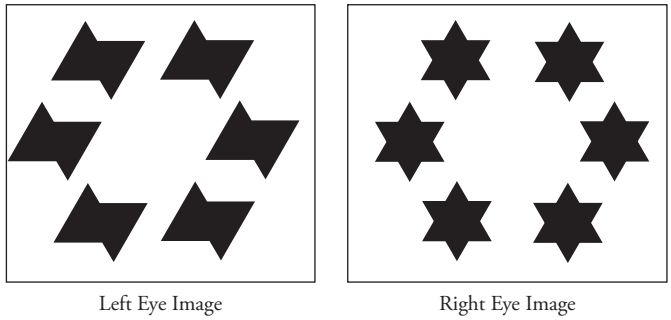


Figure VI.98-5. A stereogram by Davi Geiger and Hiroshi Ishikawa (reproduced with permission). The two images are intended to be fused. It does not matter which image is in which eye. There are no subjective contours visible monocularly but extremely strong ones emerge after fusion, resulting from the segregation of surfaces in depth.

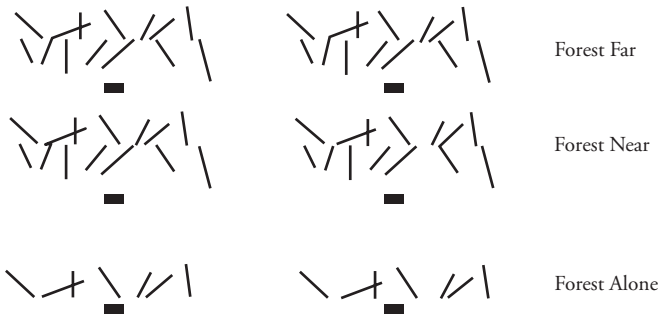


Figure VI.98-6. Three stereograms from Gillam and Nakayama (2002). Cross-fusion is assumed (left eye seeing right image and vice versa). The upper sticks in each image (called the forest) are at random stereo depths. The lower sticks (called the plane) are all at the same stereo depth. In the top stereogram the forest is more distant than the plane and a very strong subjective contour appears on fusion. In the second stereogram the forest is nearer. The subjective contour disappears. The bottom stereogram has only the forest. There is no subjective contour showing the importance of surface support in generating subjective contours. (Reproduced with permission from Gillam & Nakayama, 2002, p. 46.)

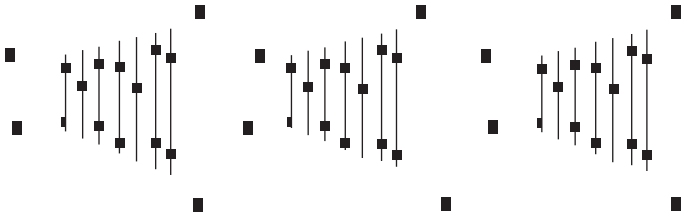


Figure VI.98-7. Stereogram from Anderson (1994, Fig. 4b). When the left and middle images are cross-fused or the middle and right images are uncrossed fused, the vertical lines appear through a phantom aperture that accounts via perceived occlusion for the differences in vertical height in the two eyes. (Reproduced from Anderson, 1994, with permission from Nature Publishing Group.)

THEORY

Physiological

Von der Heydt and Peterhans (1984) found that a number of the complex cells in monkey visual area V2 that respond to lines and edges also increased their firing rate when presented with stimuli that give rise to subjective contours in humans. These included abutting line ends and the camouflaged regions between relatable black bars spanning a black surface. The claim that these cells were responding to subjective contours was supported by the finding that their response diminished when slight changes such as small gaps were made that eliminate subjective contours in humans. These seminal findings have been the main inspiration for feed-forward or bottom-up theories of subjective contours; most notably the well-developed model by Heitger, von der Heydt, Peterhans, Rosenthaler, and Kübler (1998) in which subjective contours are "inferred" by integrating the responses of "end-stopped" cells in various alignments. "End-stopped" is a term used to describe cells whose firing rate depends not only on the orientation of a line or edge within its receptive field (the region of retina to which it responds) but also on line length. The model proposes a physiological mechanism serving a functional purpose in implementing occlusion perception. It has the advantage that it derives the direction of occlusion at a subjective contour and has been applied to occlusion relationships in photos of real scenes.

There seems to be general agreement that the small receptive fields of cells in V1 or V2 with their high spatial resolution must be involved in the generation of subjective contours. However functional magnetic resonance imaging studies (e.g., Hirsch et al., 1995) have shown that higher cortical regions, especially the lateral occipital complex are active when subjective contour producing stimuli are viewed. While the involvement of higher regions is not incompatible with feed-forward models, these findings have led to hybrid models that claim greater economy of processing than feed-forward alone. Stanley and Rubin (2003), for example, have proposed that higher visual areas, having cells with much larger receptive fields and a greater ability to pool information spatially than V2, find "salient" regions or proto surfaces and feed back their locations to cells in early cortex for the kind of detailed processing that is associated with occlusion, such as sharp contour discontinuities. There is some disagreement about exactly what higher regions may serve this function, but the general idea is a popular one. So far it accounts only for subjective contours formed around shapes. It does not account for subjective contours at abutting edges or for subjective contours at a single edge as in Figure VI.98-3c. It could, however, be postulated that salient alignments serve the same function as salient regions in demanding detailed scrutiny at lower levels for possible occlusion cues.

Physiological data associated with subjective contours is of great interest. Models based on such data are good at detecting local discontinuities, contour junctions, and global alignments. It is clear, however, that such models do not accommodate the richness or complexity of the perceptual subjective contour data. Only a few classic figures are usually considered, particularly the Kanizsa and abutting line configurations. This means not having to account for some really powerful and long-established effects. I refer particularly to the absence of subjective contours when they are predicted as in Figures VI.98-2c and VI.98-3a and the forest in front in Figure VI.98-6 and their presence when not predicted as in the interdigitated version of Figure VI.98-6. There are also huge differences that cannot be accounted for, such as Figures VI.98-3c versus VI.98-3a, especially since the latter has orthogonal contours that should favor a strong end-stopping response. One would imagine that entropy contrast should be possible to model physiologically, but it has not been. The complex effects on subjective contours of binocular manipulations such as those shown in Figure VI.98-6 cannot be accounted for by considering local neural responses to line ends, even when these are integrated. These examples show that subjective contour presence does not arise from image abutment per se but depends on both the internal depth relations of the abutting lines (plane or forest) and on their global depth relations. Such findings cannot be adequately dealt with by proposing refinement of early responses by higher processes of some kind unless the higher processes wipe out the results of the earlier processes and start again.

Nonphysiological

Several authors have tried to explain subjective contours as "cognitive"—the outcome of hypotheses accounting for "unlikely gaps" (Gregory, 1987) or probabilistic—what is seen is the most likely possibility, given the visual input and the probability of it arising from various sources. Rock (1987) proposed that we perceive the solution that avoids coincidental arrangements or alignments. For example the Kanizsa figure could be produced by a triangle occluding three circles or by three notched circles arranged so that the notches become multiply aligned—a highly coincidental possibility. We prefer seeing the occluder to the multiple coincidences. This is related to another popular theory of perceptual organization: the generic versus accidental view theory. This proposes that we see from among the possible outcomes the one that would arise from a number of different viewpoints or vantage points rather than one that would only arise from a single vantage point.

All of these approaches one way or another postulate the avoidance of improbable solutions. They resemble the more formal Bayesian approach in which perception is based on optimizing probabilistic inference. The flaw in such theories is not the proposal that the visual system incorporates ecologically probable outcomes in its mechanisms and processes. It is in assuming that such entailments act as "laws" or rules in which the outcome can always be predicted on a probabilistic basis. This is clearly not the case. First, it is highly implausible that there is an online calculation of possible outcomes from which the most probable or the most simple is chosen. Can it then be assumed that whatever mechanisms exist always give the most probable result? Mechanisms have presumably evolved or been learned in ecological settings to deal with occlusion relationships, but it is necessary to determine empirically what these mechanisms are and on what information they depend. A subjective contour is not always seen when occlusion at that location would be the most likely outcome or the one that would best avoid coincidence. For example when the gaps between notched circles in the Kanizsa figure are increased, the subjective contour diminishes in strength even though an occluding triangle surely remains the most probable solution and the one that best avoids coincidences. This effect may well have a neural basis. Likewise, in the binocular figures shown in Figure VI.98-6 the forest alone does not produce a subjective contour even though again it is highly unlikely that a set of lines at different depths would be aligned along one edge except by an occluder. Perceiving this alignment without occlusion would be a highly "accidental" view. Yet that is what is seen unless the nearer plane is present.

These examples point to a common requirement of seeing subjective contours. In addition to inducers that would seem to be very likely to arise from an occlusion, there usually has to be support for the presence of a surface in a plausible occluding location. The visual system requires a certain amount of evidence to support the percept, and this evidence must be discovered empirically.

We conclude that in explaining subjective contours there is no substitute for exploring the conditions that give rise to them as a way of determining the visual mechanisms responsible. There is no principle that can account for all cases. The study of this phenomenon is an excellent way of studying the perception of occlusion, especially its global determinants.

REFERENCES

Albert, M. K., & Hoffman, D. D. (2000). The generic-viewpoint assumption and illusory contours. *Perception*, *29*, 303–312.

Anderson, B. L. (1994). The role of partial occlusion in stereopsis. *Nature*, *367*, 365–367.

Coren, S. (1972). Subjective contours and apparent depth. *Psychological Review, 79*, 359–367.

Ehrenstein, W. (1941). Uber Abwandlundgen der L. Hermannschen Helliskeitserscheinung. *Zeitschrift für Psychologie. 150*, 83–91. (English translation: A. Hogg, Modifications of the brightness phenomenon of L. Hermann, in S. Petry & G. E. Meyer [Eds.], *The perception of illusory contours*. New York: Springer-Verlag, 1987)

Gillam, B., & Chan, W.M. (2002). Grouping has a negative effect on both subjective contours and perceived occlusion at T-junctions. *Psychological Science*, *13*, 279–283.

Gillam, B., & Nakayama, K. (2002). Subjective contours at line terminations depend on scene layout analysis, not image processing. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 43–53.

Gregory, R. L. (1987). Illusory contours and occluding surfaces. In S. Petry & G. E. Meyer (Eds.), *The perception of illusory contours* (pp. 81–89). New York: Springer-Verlag.

Heitger, F., von der Heydt, R., Peterhans, E., Rosenthaler, L., & Kübler, O. (1998). Simulation of neural contour mechanisms: Representing anomalous contours. *Image and Vision Computing, 16*, 407–421.

Hirsch, J., DeLaPaz, R. L., Relkin, N. R., Victor, J., Kim, K., Li, T., … Shapley R. (1995). Illusory contours activate specific regions in human visual cortex: Evidence from functional magnetic resonance imaging. *Proceedings of the National Academy of Sciences USA, 92*(14), 6469–6473.

Kanizsa, G. (1955). Quasi-percettivi in campi con stimolazione omogenea. *Rivista di Psicologia, 49*, 7–30. (English translation: W. Gerbino, Quasi-perceptual margins in homogeneously stimulated fields, in S. Petry & G. E. Meyer [Eds.], *The perception of illusory contours*. New York: Springer-Verlag, 1987)

Kanizsa, G. (1979). *Organization in vision*. New York: Praeger.

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, *23*, 141–221.

Kennedy, J. M. (1978). Illusory contours at the ends of lines. *Perception, 7*, 605–607.

Nakayama, K., & Shimojo, S. (1990). Da Vinci stereopsis: Depth and subjective occluding contours from unpaired image points. *Vision Research*, *30*, 1811–1825.

Rock, I. (1987). A problem solving approach to illusory contours. In S. Petry & G. E. Meyer (Eds.), *The perception of illusory contours* (pp. 62–69). New York: Springer-Verlag.

Schumann, F. (1904). Beiträge zur Analyse der Gesichtswahrnehmungen. *Zeitschrift fur Psychologie, 23*, 1–32.

Stanley, D. A., & Rubin, N. (2003). fMRI response to illusory contours and salient regions. *Neuron*, *37*, 323–331.

Szily, A. von (1921). Stereoskopische versuche mit Schattenrissen. *Albrecht von Graefes Archiv für Ophthalmologie, 105*, 964–972. (English translation: W. Ehrenstein & B. Gillam [1999], Early demonstrations of subjective contours, amodal completion and half-occlusions. *Perception 27*, 1407–1416.

von der Heydt, R., Peterhans, E., & Baumgartner, G. (1984). Illusory contours and cortical neural responses. *Science*, *224*, 1260–1262.

Chapter 99
The Ternus Effect

Elisabeth Hein

DESCRIPTION OF THE TERNUS EFFECT

The Ternus effect (sometimes also referred to as the Ternus-Pikler effect) is an illusion that usually consists of three discs that are presented in a row. From one frame to the next, the set of stimuli is shifted by one position to the left or right, such that the center stimuli are presented at the same position across frames and the first stimulus changes position (Fig. VI.99-1a). These two stimulus frames are commonly presented as cycling back and forth but sometimes also presented as just once each. This display can be perceived in (at least) two very different ways (Video VI.99-1): as all stimuli moving together as a group (group motion) or as the first of the stimuli jumping across the other two, the center stimuli, while they seem not to move at all (element motion). Which percept is seen depends on many different factors, for example the interstimulus interval (ISI) between the two stimulus frames or the similarity of the Ternus elements (as detailed below).

A compelling variation of the Ternus effect was introduced by Boi, Öğmen, Krummenacher, Otto, and Herzog (2009), who added little points to each of the classic Ternus elements. These points can be perceived either as rotating (counter-)clockwise or as randomly changing position on the Ternus elements. Which of these alternatives is seen depends on whether the Ternus elements themselves are organized in terms of group or element motion (Video VI.99-2).

SHORT HISTORY OF THE TERNUS EFFECT AND RELEVANCE FOR MODERN VISION RESEARCH

Julius Pikler first described the stimulus configuration used in the Ternus effect in 1917 as one of many different apparent motion configurations that he tested. With the stimulus parameters he used, he saw the stimuli—in a two-disc version of the Ternus display—moving together shifting by one position, or what we call group motion today. He found it remarkable that despite the center stimulus being presented at the exact same position, it was not perceived as staying at this position but rather as moving to another position. This fit well within his accommodation theory (termed *Anpassungstheorie* by Ternus) of visual perception that—contrary to what many believed at the time—there were perceptions that are not based directly on the sensory activation but on the interpretation or spontaneous ad hoc accommodation of the visual system to the stimulus that goes beyond the pure physical stimulation, but should neither be dependent on prior knowledge or experience.

A couple of years later, Joseph Ternus (1926) used the display to investigate the conditions under which we perceive object identity (*phänomenale Identität*). He as well was interested in the Ternus display as an important exception to the general principle that a stimulus presented at the same location over time should always be perceived as the same stimulus; that is, object identity should be preserved. He investigated the Ternus display more closely and found conditions in which the center stimuli either changed position or not (i.e., the percept had been organized in terms of group or element motion). He used, for example, two lines of different length as Ternus stimuli and showed that, in this case, one could perceive the first element jumping across the other element (i.e., element motion). On the other hand, when both stimuli were identical, they appeared to both move by one position (i.e., group motion). He also discovered that not only the line length but also the distance between the Ternus elements and the presentation duration of the Ternus elements could induce element and group motion depending on the conditions. He reported, for example, that the longer the distance between the elements, especially when they are small, the more likely one will perceive element motion. Based on these and related demonstrations, Ternus formulated the thesis that the perception of object identity is determined by the whole Gestalt and not the parts presented in each frame. Accordingly, he argued that the element motion percept would be the special case, not the group motion percept.

Fifty years later Pantle and Picciano (1976) revived the Ternus display, emphasizing the bistable nature of the display and highlighting its potential for investigating (apparent) motion perception. Since then the Ternus display has been studied extensively and is still used to investigate fundamental research questions on how the visual system interprets dynamic stimuli and computes apparent motion and correspondence (i.e., the question of how the visual system is able to maintain the identity of stimuli over time and across changes in location, occlusion of the stimuli, or eye and body movements of the observer; Dawson, 1991; Ullman, 1979). The idea being that, depending on which Ternus motion is

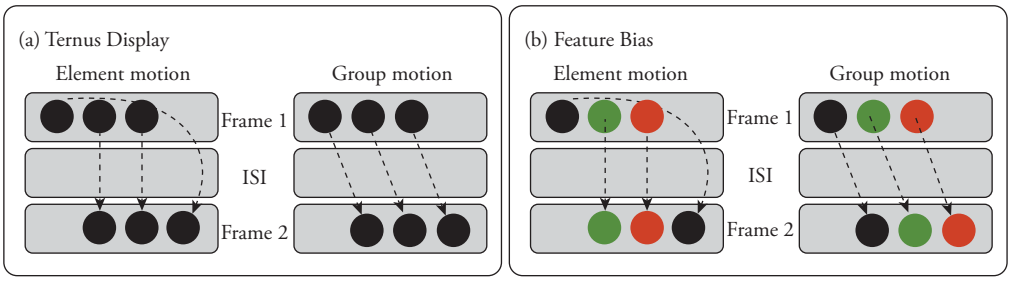
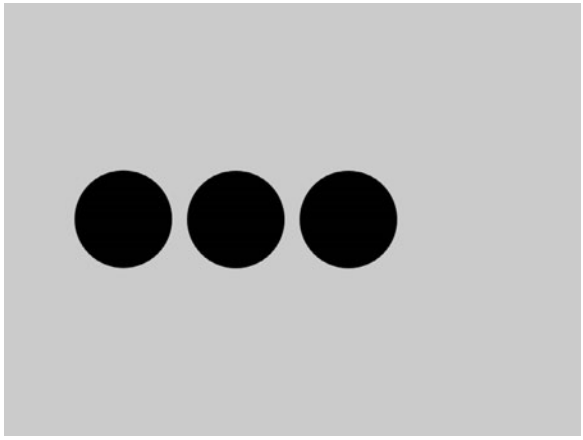
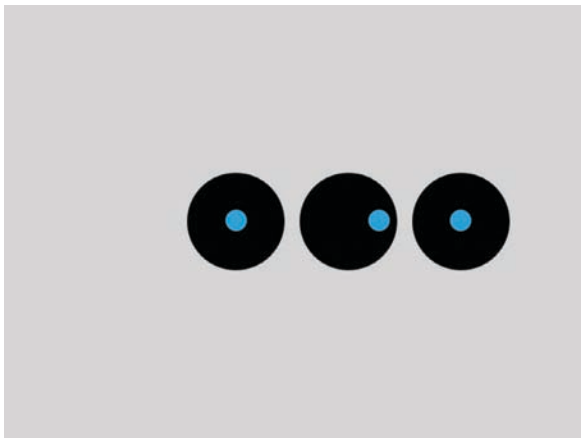


Figure VI.99-1. (a) Basic configuration of stimuli used for a three-disc Ternus display and different percepts (element and group motion; dashed lines). (b) Feature bias based on color. If the feature bias is compatible with element motion (left part of [b]), element motion is more likely to be perceived. If the feature bias is compatible with group motion (right part of [b]), more group motion is perceived. The dashed lines in figure (a) and (b) are only for illustration purposes and are not presented in the display.



Video VI.99-1. Illustration of two possible percepts in the Ternus display, an ambiguous apparent motion display, in which two frames of three stimuli separated by an empty interstimulus frame are alternating with each other (as in Fig. VI.99-1a). The three Ternus elements can be perceived as moving together as a group (group motion) or as the first of the elements jumping across the other two stationary central elements (element motion). The video uses different interstimulus intervals (ISI) between the two stimulus frames in order to bias the different motion percepts. (Figure by Elisabeth Hein.)



Video VI.99-2. Illustration of a variation of the original Ternus display introduced by Boi, Öğmen, Krummenacher, Otto, and Herzog (2009). The little points on top of the Ternus elements can be either perceived as rotating (counter-)clockwise or as randomly changing position on the Ternus elements. Different ISI are used in order to change the correspondence solution established between the Ternus elements and thus the percept of the little points. (Figure by Elisabeth Hein.)

perceived, correspondence is established between different discs. So one can make changes to the Ternus display and assess how these changes affect correspondence by recording if observers perceive element or group motion. Following this logic, the Ternus display has been used to examine a long-lasting debate in the literature about whether features, like luminance or color, can affect the correspondence solution (e.g., Dawson, Nevin-Medows, & Wright, 1994; Hein & Moore, 2012; Petersik & Rice, 2008).

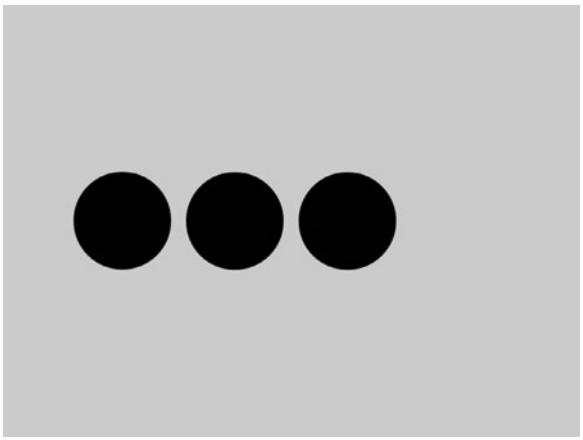
Furthermore, the Ternus display has been used more recently as a tool to investigate the fundamental question of the reference frame through which stimuli are presented by the visual system—that is, whether or not they are processed in retinotopic coordinates (Boi et al., 2009; Boi, Öğmen, & Herzog, 2011; Boi, Vergeer, Öğmen, & Herzog, 2011; Öğmen, Otto, & Herzog, 2006; Thunell, Plomp, Ögmen, & Herzog, 2015; van Boxtel & Koch, 2012). "Retinotopic" means that a stimulus location is encoded in terms of the position the stimulus projects to on the retina. This implies that every time we move our eyes, the position of the stimulus on the retina changes, even though the stimulus does not move in the "real world." "Nonretinotopic," on the other hand, means that the visual system has computed the location of a stimulus in an external reference frame, or in terms of where the stimulus is located in the world, compensating for the location changes on the retina. The Ternus display perceived in group motion allows us to distinguish between retinotopic and nonretinotopic processing. This is because the central Ternus elements in the first frame and the first elements in the second frame occupy the same retinotopic position (at least as long fixation is maintained). The central Ternus element in the first frame, however, is matched nonretinotopically with the central element in the second frame. By presenting different stimuli on top of the Ternus elements, one can then examine whether perception matches the retinotopic or the nonretinotopic stimulus conditions. Using this technique Boi et al. (2009) and Boi, Öğmen, et al. (2011) found, for example, that motion aftereffects occur largely in retinotopic coordinates while visual search seems to occur rather in nonretinotopic coordinates.

The Ternus display has also been used for examining how different sensory modalities, like vision, audition, and touch, interact with each other (Chen, Shi, & Müller, 2010; Harrar & Harris, 2007; Shi, Chen, & Müller, 2010). This research has shown, for example, that there are tactile and visuotactile versions of the Ternus effect that are very

similar to the visual Ternus (Harrar & Harris, 2007) and that auditory stimuli presented before and after or synchronously with the visual Ternus display can affect Ternus perception (Shi et al., 2010).

FACTORS AFFECTING THE TERNUS EFFECT

We know today that the Ternus display can be influenced by many different factors. These can be divided into two broad categories: "low-level" factors that are defined directly in the image (e.g., the luminance or retinal size of a stimulus) and "higher-level" factors that are extracted from the basic image features following further processing of the visual system (e.g., perceptual grouping or perceived size). The factors that can affect the Ternus display have interested researchers because they can shed light on the underlying mechanisms of the Ternus display in particular and motion perception and correspondence in general (see later discussion).



Video VI.99-3. Illustration of a feature bias based on color: the colors of the Ternus elements can be compatible with element motion (as in Fig. VI.99-1b, left display), in which case it is more likely that element motion will be perceived. Alternatively, the colors of the Ternus elements can be compatible with group motion (as in Fig. VI.99-1b, right display), in which case group motion will be more likely perceived. (Figure by Elisabeth Hein.)

Low-Level Factors

One important low-level group of factors that can influence the perception of the Ternus display is spatiotemporal factors (i.e., the temporal and spatial distance between the Ternus elements). An example is the ISI, or the time between the two stimulus frames. The longer the ISI, the more likely observers will perceive group motion (Pantle & Picciano, 1976; Petersik & Pantle, 1979; see Video VI.99-1). Although the exact ISI at which the percept changes from element to group motion varies across stimulus conditions and observers, it is usually somewhere around 40 ms. Another temporal parameter that influences how the Ternus display is perceived is how long each Ternus frame is presented (Breitmeyer & Ritter 1986a; Petersik & Pantle, 1979; Ternus, 1926): the longer the frame duration, the more likely observers will see group motion. In terms of spatial parameters, the separation or spatial distance between the Ternus elements can affect how the display is perceived: Specifically, the larger the spacing between elements within a frame, the less likely group motion will be perceived (Alais & Lorenceau, 2002; Pantle & Petersik, 1980; Ternus, 1926). Furthermore, the eccentricity at which the stimuli are seen can influence perception. In general, the farther the Ternus display is presented away from fixation (i.e., the more it is presented in the periphery), the more group motion is perceived (Aydin, Herzog, & Öğmen, 2011; Breitmeyer & Ritter, 1986b; but see also Alais & Lorenceau, 2002, and Petersik, 2009, who found different effects on eccentricity).

In addition to spatiotemporal factors, another group of low-level factors that strongly affects the perception of the Ternus display is the appearance of the individual elements or their features, such as the luminance/contrast, spatial frequency, orientation, color, or size of the elements (Alais & Lorenceau, 2002; Breitmeyer & Ritter 1986a; Casco, 1990; Dawson et al., 1994; Hein & Moore, 2012; Hein & Cavanagh, 2012; Kramer & Rudd, 1999; Kramer & Yantis, 1997; Petersik & Rice, 2008; Ternus, 1926; Wallace & Scott-Samuel, 2007). In general, the more distinct the elements are from each other, the more likely element motion will be perceived. Even stronger is the effect of a feature-based bias: if the features of the elements are constructed in such a way that they are compatible with element motion, it is very likely that element motion will be perceived. Alternatively, if the features are compatible with group motion, it is very likely that group motion will be perceived (Figure VI.99-1b; Video VI.99-3).

Higher-Level Factors

Both spatiotemporal and feature parameters can be considered low level in that they are defined directly in the image. More recent studies have shown that, contrary to what has been believed for a long time, other, more high-level factors can also influence the perception of the Ternus display. These factors include the perceived organization of a scene (He & Ooi, 1999; Hein & Moore, 2014; Kramer & Yantis, 1997; Ma-Wyatt, Clifford, & Wenderoth, 2005; Ternus, 1926; Yantis, 1995), spatial attention (Aydin et al., 2011; Ternus, 1926), visual working memory (Hein, Moore, & Hollingworth, 2013), and even lexical information (Chen & Zhou, 2011) and the semantic context (Yu, 2000; Hsu, Taylor & Pratt, 2015). For example, to examine the effect of the perceived organization of the scene, He and Ooi showed that if the center elements are perceptually grouped away from the first element by two lines of vertical elements (Fig. VI.99-2a), different positions in depth, or different surfaces, less group motion is perceived. As another example of the importance of the perceived rather than low-level retinal factors, Hein and Moore showed that the perceived and not the retinal size of the Ternus element influences perception. They presented an occluder with a small hole in front of the Ternus display, thus reducing the retinal size of the Ternus disc, while observers perceived a large disc that extended behind the small occluder hole (Fig. VI.99-2b). The Ternus percept matched the perceived and not the retinal stimulus size; that is, participants reported seeing predominantly group and not element motion in this condition.

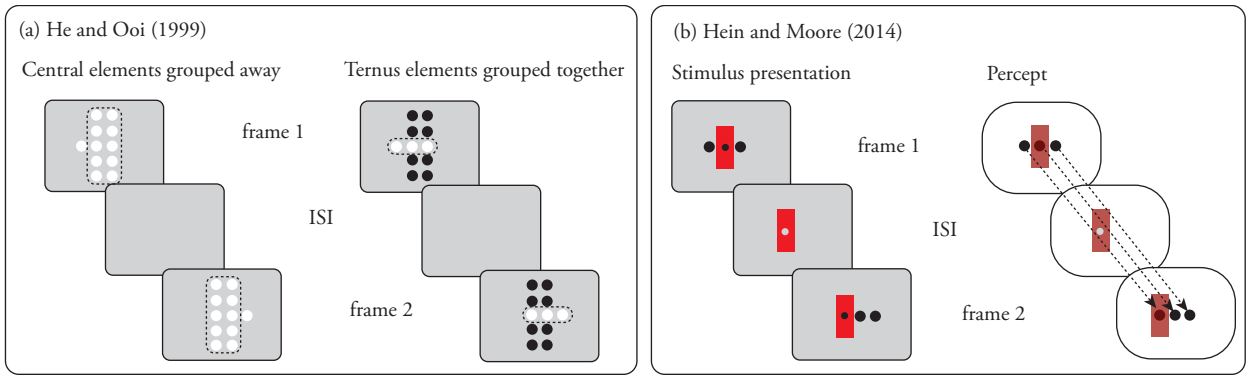


Figure VI.99-2. (a) Example of stimulus arrangement used by He and Ooi (1999): The center Ternus elements could be grouped away from the first element by two lines of vertical elements that matched the identity of the center elements (left part of [a]), which resulted in more element motion being perceived. The center elements could also be grouped together with the first element by two lines of vertical elements that did not match the identity of the central elements (right part of [a]), in which case more group motion was perceived. The dashed lines are for illustration purposes only and are not presented in the displays. (b) Stimulus arrangement used by Hein and Moore (2014) and perception of Ternus motion. The dashed arrows illustrate the group motion percept.

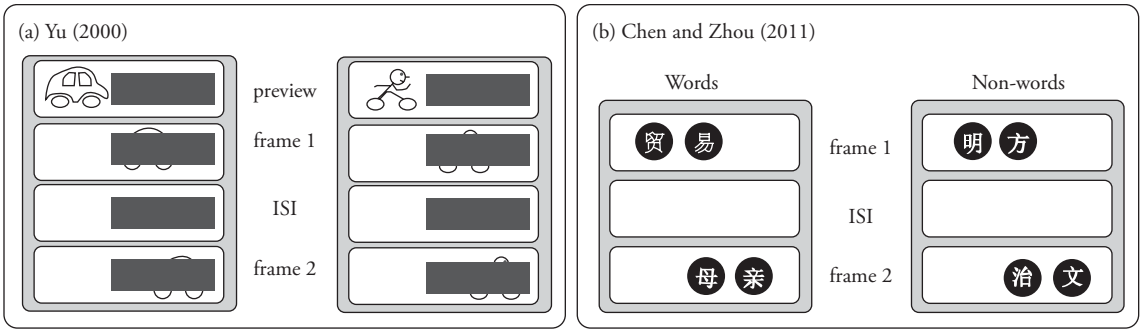


Figure VI.99-3. (a) Illustration of stimuli used by Yu (2000) to examine if semantic information (car vs. person) can affect Ternus perception. (b) Illustration of word and non-word stimuli used by Chen and Zhou (2011) to investigate the influence of lexical information (words vs. non-words) on the Ternus percept.

In addition to the perceived organization of the scene, even more high-level processes like attention and memory can influence the perception of the Ternus display. Aydin et al. (2011), for example, manipulated attention by introducing a second attention-demanding task at the fovea and presenting the Ternus display in the periphery. Participants perceived less group motion when their attention was withdrawn from the Ternus display with this dual task, suggesting that the group motion percept requires more attentional resources. Regarding the influence of memory, Hein et al. (2013) showed that a color held in visual working memory could influence Ternus motion in an ambiguous Ternus display in which different isoluminant colors introduced an element and a group bias within the same display. The color bias that matched the color held in memory had a stronger effect on the Ternus motion; that is, when the memory color matched element bias color, more element motion was seen and when the memory color matched the group bias color, more group motion was seen.

Finally, even as high-level processes such as semantic and lexical processing can influence Ternus perception. Yu (2000) demonstrated that the semantic context could influence what motion is perceived in the Ternus display: that data showed a slight yet significant decrease in group motion reports when the previous context suggested that the Ternus discs represented the two feet of a person compared to a context in which the Ternus discs represented the two tires of a car (Fig. VI.99-3a; for an even larger semantic effect with frogs as Ternus elements see Hsu et al., 2015). Using Chinese characters presented on top of the Ternus discs, Chen and Zhou (2011) demonstrated a lexical effect on the Ternus: if the characters presented on the Ternus discs of one frame formed a compound word, participants were more likely to perceive group motion than when the two characters did not form a word (Fig. VI.99-3b). These results suggest that lexical information can affect the Ternus percept, even though it is task-irrelevant.

THEORIES OF THE TERNUS EFFECT

A variety of theories have been developed during the 100 years since the illusion was first described. These theories can be categorized broadly into motion-based theories and object-based theories. Most motion-based theories emphasize the role of basic motion detectors or spatiotemporal filters as they are used in classic motion theories (e.g., Adelson & Bergen, 1985; Pooresmaeili, Cicchini, Morrone, & Burr, 2012; Werkhoven, Sperling, & Chubb, 1993; but see also Clarke, Repnow, Öğmen, & Herzog,

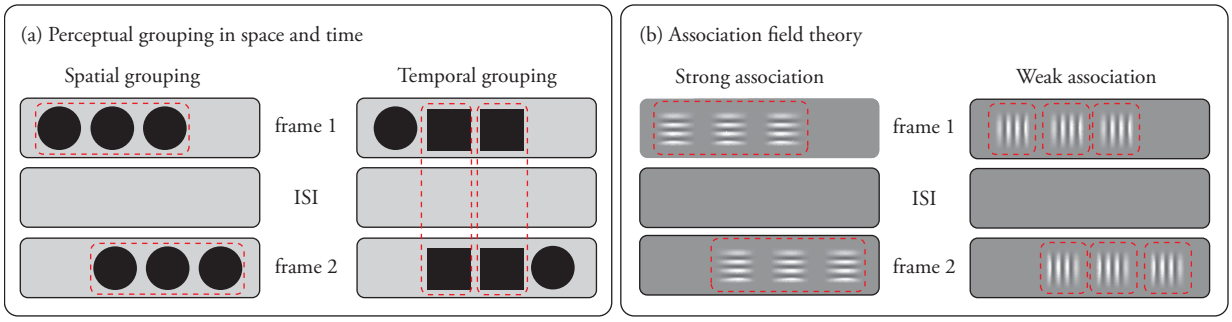


Figure VI.99-4. (a) Illustration of spatial and temporal grouping as suggested by Kramer and Yantis (1997). For spatial grouping, the similarity of elements within each Ternus frame is important (left part of [a]), and the higher the similarity, the more group motion is perceived. For temporal grouping the similarity of elements across Ternus frames is important (right part of [a]), and the higher the similarity of these elements, the more element motion is perceived. (b) Illustration of the association field theory by Alais and Lorenceau (2002). If the elements are collinear (left part of [b]), more group motion is perceived due to stronger lateral interactions; otherwise more element motion is perceived (right part of [b]).

2013). These motion detectors determine the direction of the main motion energy between stimuli based on the changes in luminance contrast at the different stimulus positions over time, which in turn determine the motion percept (van Santen & Sperling, 1985). In contrast, object-based theories highlight the perceptual organization or the relations between the different elements of the display and motion is perceived as a consequence of which objects are more likely to belong together (e.g., based on Gestalt grouping principles; He & Ooi, 1999; Hein & Moore, 2014; Kramer & Rudd, 1999; Öğmen, 2007; Petersik & Rice, 2008; Ternus, 1926; Wallace & Scott-Samuel, 2007). We next present examples of each type of theory in more detail (see also Petersik & Rice, 2006, for a recent review of explanations of the Ternus effect).

Motion-Based Theories

A classic example of a motion-based theory for the Ternus display is based on the distinction between short- versus long-range motion processes (Braddick, 1980). Short-range processes are supposed to work over small stimulus displacements and short temporal intervals, while long-range processes operate over longer spatial displacements and time intervals. It has been suggested that the short-term motion process could detect the absence of motion of the central Ternus elements and therefore would be responsible for the perception of element motion, while the long-range motion processes would be responsible for the perception of group motion (Braddick & Adlard, 1978; Pantle & Petersik, 1980). This account has been criticized, however, based on the two-motion process distinction in itself (Cavanagh, 1992; Cavanagh & Mather, 1989), as well as its usefulness to explain the Ternus percepts (Odic & Pratt, 2008, 2010; Scott-Samuel & Hess, 2001).

Several other motion-based theories have been developed that do not depend on the two-motion process distinction. A recent example of such a theory is the *differential-activation theory* adapted to the Ternus display by Odic and Pratt (2008; but see Petersik, 2010; Silva & Petersik, 2015). Following this theory, it is sufficient to assume that the perception of the Ternus display is determined by the change in luminance across the different Ternus frames. If the difference in luminance change is large enough, an element is perceived as moving to its nearest neighbor available (i.e., the *nearest neighbor principle*; Burt & Sperling, 1981; Dawson, 1991). The first element always has a high change in luminance, as it disappears from one frame to the next, and is therefore always perceived as moving. Where this element moves, however, depends on which element is the next possible candidate. If the central elements are perceived as stationary, because the luminance change is too small, the first element moves to the last element in the second frame, as it is the next available (i.e., not perceived as stationary). Element motion is perceived in this case. On the other hand, if the central elements are not perceived as stationary, the next available neighbor is the next element in the row; thus each element is perceived to move by one position, and group motion is perceived. This theory is very similar to the *pattern persistence theory* proposed by Breitmeyer and Ritter (1986a, 1986b; see also Grossberg & Rudd, 1992). In particular, they proposed that pattern persistence could be responsible for signaling the stationary nature of the central elements, in a way closing the gap between successive stimulus presentations of the central elements and thus inhibiting the perception of group motion. If, on the other hand, the gap is too long to be closed by pattern persistence, then motion is perceived to the nearest neighbor; that is, all the elements move by one position.

Object-Based Theories

In addition to these motion-based theories, several object-based theories of the Ternus display have also been proposed. One example is the *perceptual grouping in space and time theory* by Kramer and Yantis (1997; Kramer & Rudd, 1999; see also He & Ooi, 1999; Petersik & Rice, 2008; Wallace & Scott-Samuel, 2007). Kramer and colleagues suggested that the perception of the Ternus display depends on the competition between spatial and temporal grouping (Fig. VI.99-4a). The amount of spatial grouping is determined by factors that affect how much the elements within a single Ternus frame are perceived as one coherent group (i.e., the similarity of the elements and their general

configuration or the context). The strength of temporal grouping, on the other hand, is determined by the temporal proximity and the similarity of individual elements across successive frames. The stronger the spatial grouping and the weaker the temporal, the more group motion and the less element motion is seen and vice versa.

Another idea that is in some ways similar to Kramer and Yantis's (1997; Kramer & Rudd, 1999) spatial grouping has been suggested by Alais and Lorenceau (2002). They proposed that the perception of the Ternus display depends on the strength of an *association field* that determines how strongly neighboring stimuli are associated with each other, with this association being influenced by colinearity and distance of the elements. The stronger the association field, the more likely group motion will be perceived. (Fig. VI.99-4b).

Finally, Hein and Cavanagh (2012; see also Hein & Moore, 2012; Hein & Moore, 2014) suggested an object-based theory to explain correspondence in the Ternus display based on the *attentional pointer account* (Cavanagh, 1992). In contrast to the other object-based theories this account does not rely on the grouping or association strengh of all the elements together, but rather on the perceived similarity of the individual elements, pair by pair. According to the theory of Hein et al. the motion percept in the Ternus display is determined by attentional pointers that establish correspondence over space and time on the basis of the individual perceived object representations of each Ternus element that are created from the further processed image features, as for example perceived size, that exist in a non-retinotopic reference frame and that may include the semantic and lexical information associated with the objects. Motion is then assigned as a second step between the most similar object representations (see also Ramachandran, Armel, Foster, & Stoddard, 1998).

Conclusions

Motion-based theories can explain the influence on the Ternus display of a number of low-level parameters that are defined directly in the image, for example effects of ISI or luminance. Object-based theories, on the other hand, can explain the influence of more high-level factors, for example the perceived organization of a scene or the semantic content, that do not vary at the image level and are therefore unlikely to affect processes suggested by motion-based theories. It is very likely that lower-level motion-based accounts as well as higher-level object-based accounts are necessary to explain the variety of different factors that have been shown to affect perception of the Ternus display. Thus the different accounts of the Ternus display are not mutually exclusive, and both types of mechanism may be involved. In particular, they might work together or separately depending on the available information (i.e., the stimulus parameters or the complexity of the correspondence problem posed).

REFERENCES

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, *2*(2), 284–299.

Alais, D., & Lorenceau, J. (2002). Perceptual grouping in the Ternus display: Evidence for an "association field" in apparent motion. *Vision Research*, *42*(8), 1005–1016.

Aydin, M., Herzog, M., & Öğmen, H. (2011). Attention modulates spatio-temporal grouping. *Vision Research*, *51*(4), 435–446. doi:10.1016/j.visres.2010.12.013

Boi, M., Öğmen, H., & Herzog, M. (2011). Motion and tilt aftereffects occur largely in retinal, not in object, coordinates in the Ternus-Pikler display. *Journal of Vision*, *11*(3):7, 1–11.

Boi, M., Öğmen, H., Krummenacher, J., Otto, T., & Herzog, M. (2009). A (fascinating) litmus test for human retino- vs. non-retinotopic processing. *Journal of Vision*, *9*(13):5, 1–11.

Boi, M., Vergeer, M., Öğmen, H., & Herzog, M. (2011). Nonretinotopic exogenous attention. *Current Biology, 21*(20), 1732–1737.

Braddick, O. J. (1980). Low-level and high-level processes in apparent motion. *Philosophical Transactions of the Royal Society of London B*, *290*, 137–151.

Braddick, O. J., & Adlard, A. (1978). Apparent motion and the motion detector. In J. Armington, J. Krauskopf, & B. R. Wooten (Eds.), *Visual psychophysics and physiology* (pp. 417–426). New York: Academic Press.

Breitmeyer, B., & Ritter, A. (1986a). The role of visual pattern persistence in bistable stroboscopic motion. *Vision Research*, *26*(11), 1801–1806.

Breitmeyer, B. G., & Ritter, A. (1986b). Visual persistence and the effect of eccentric viewing, element size, and frame duraton on bistable stroboscopic motion percepts. *Perception & Psychophysics*, *39*(4), 275–280.

Burt, P., & Sperling, G. (1981). Time, distance, and feature trade-offs in visual apparent motion. *Psychological Review*, *88*(2), 171–195.

Casco, C. (1990). The relationship between visual persistence and event perception in bistable motion display. *Perception*, *19*(4), 437–445.

Cavanagh, P. (1992). Attention-based motion perception. *Science*, *257*(5076), 1563–1565.

Cavanagh, P., & Mather, G. (1989). Motion: The long and short of it. *Spatial Vision*, *4*(2–3), 103–129.

Chen, L., Shi, Z., & Müller, H. (2010). Influences of intra- and crossmodal grouping on visual and tactile Ternus apparent motion. *Brain Research*, *1354*, 152–162.

Chen, L., & Zhou, X. (2011). Visual apparent motion can be modulated by task-irrelevant lexical information. *Attention, Perception, & Psychophysics*, *73*(4), 1010–1015.

Clarke, A., Repnow, M., Öğmen, H., & Herzog, M. (2013). Does spatio-temporal filtering account for nonretinotopic motion perception? Comment on Pooresmaeili, Cicchini, Morrone, and Burr (2012). *Journal of Vision*, *13*(10): 19, 1–14. doi:10.1167/13.10.19

Dawson, M. (1991). The how and why of what went where in apparent motion: Modeling solutions to the motion correspondence problem. *Psychological Review*, *98*(4), 569–603.

Dawson, M., Nevin-Meadows, N., & Wright, R. (1994). Polarity matching in the Ternus configuration. *Vision Research*, *34*(24), 3347–3359.

Grossberg, S., & Rudd, M. (1992). Cortical dynamics of visual motion perception: Short-range and long-range apparent motion. *Psychological Review*, *99*(1), 78–121.

Harrar, V., & Harris, L. (2007). Multimodal Ternus: Visual, tactile, and visuo-tactile grouping in apparent motion. *Perception*, *36*(10), 1455–1464.

He, Z., & Ooi, T. (1999). Perceptual organization of apparent motion in the Ternus display. *Perception*, *28*(7), 877–92.

Hein, E., & Cavanagh, P. (2012). Motion correspondence in the Ternus display shows feature bias in spatiotopic coordinates. *Journal of Vision*, *12*(7): 16, 1–14.

Hein, E., & Moore, C. M. (2012). Spatio-temporal priority revisited: The role of feature identity and similarity for object

correspondence in apparent motion. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(4), 975–988.
Hein, E., & Moore, C. M. (2014). Evidence for scene-based motion correspondence. *Attention, Perception & Psychophysics*, *76*(3), 793–804.
Hein, E., Moore, C. M., & Hollingworth, A. (2013). The influence of visual working memory on correspondence in the Ternus display. *Journal of Vision, 13*(9), 61.
Hsu, P., Taylor, J. T., & Pratt, J. (2015). Frogs jump forward: Semantic knowledge influences the perception of element motion in the Ternus display. *Perception*, *44*(7), 779-789.
Kramer, P., & Rudd, M. (1999). Visible persistence and form correspondence in Ternus apparent motion. *Perception & Psychophysics*, *61*(5), 952–962.
Kramer, P., & Yantis, S. (1997). Perceptual grouping in space and time: Evidence from the Ternus display. *Perception & Psychophysics*, *59*(1), 87–99.
Ma-Wyatt, A., Clifford, C., & Wenderoth, P. (2005). Contrast configuration influences grouping in apparent motion. *Perception*, *34*(6), 669–685.
Odic, D., & Pratt, J. (2008). Solving the correspondence problem within the Ternus display: The differential-activation theory. *Perception*, *37*(12), 1790–1804.
Odic, D., & Pratt, J. (2010). Differential-activation theory can account for the Ternus display: Rejoinder to Petersik. *Perception*, *39*(5), 711–717.
Öğmen, H. (2007). A theory of moving form perception: Synergy between masking, perceptual grouping, and motion computation in retinotopic and non-retinotopic representations. *Advances in Cognitive Psychology*, *3*(1–2), 67–84.
Öğmen, H., Otto, T., & Herzog, M. (2006). Perceptual grouping induces non-retinotopic feature attribution in human vision. *Vision Research*, *46*(19), 3234–3242. doi:10.1016/j.visres.2006.04.007
Pantle, A., & Petersik, J. (1980). Effects of spatial parameters on the perceptual organization of a bistable motion display. *Perception & Psychophysics*, *27*(4), 307–312.
Pantle, A., & Picciano, L. (1976). A multistable movement display: Evidence for two separate motion systems in human vision. *Science*, *193*(4252), 500–502.
Petersik, J. (2009). Orientation anisotropy in the Ternus phenomenon. *Perceptual and Motor Skills*, *108*, 405–410.
Petersik, J. (2010). Ternus effect: Two processes or differential activation? Comments on Odic and Pratt's 2008 paper. *Perception*, *39*, 705–710.
Petersik, J., & Pantle, A. (1979). Factors controlling the competing sensations produced by a bistable stroboscopic motion display. *Vision Research*, *19*(2), 143–154.
Petersik, J., & Rice, C. (2006). The evolution of explanations of a perceptual phenomenon: A case history using the Ternus effect. *Perception*, *35*(6), 807–821.
Petersik, J., & Rice, C. (2008). Spatial correspondence and relation correspondence: Grouping factors that influence perception of the Ternus display. *Perception*, *37*(5), 725–739.
Pikler, J. (1917). *Sinnesphysiologische Untersuchungen*. Leipzig: Barth.
Pooresmaeili, A., Cicchini, G., Morrone, M., & Burr, D. (2012). "Non-retinotopic processing" in Ternus motion displays modeled by spatiotemporal filters. *Journal of Vision*, *12*(1), 10, 1–15.
Ramachandran, V. S., Armel, C., Foster, C., & Stoddard, R. (1998). Object recognition can drive motion perception. *Nature*, *395*(6705), 852–853.
Scott-Samuel, N., & Hess, R. (2001). What does the Ternus display tell us about motion processing in human vision? *Perception*, *30*(10), 1179–1188.
Shi, Z., Chen, L., & Müller, H. (2010). Auditory temporal modulation of the visual Ternus effect: The influence of time interval. *Experimental Brain Research*, *203*(4), 723–735.
Silva, M. R., & Petersik, J. T. (2015). Flicker-defined forms in the Ternus display. *Perception*, *44*(3), 315-327.
Ternus, J. (1926). Experimentelle Untersuchungen über phänomenale Identität. *Psychologische Forschung, 7*, 81–136. (English translation in W. D. Ellis [Ed.], *A sourcebook of Gestalt psychology*. New York: Humanities Press, 1950)
Thunell, E., Plomp, G., Ögmen, H., & Herzog, M. H. (2015). Eeg correlates of relative motion encoding. *Brain Topography*, doi:10.1007/s10548-015-04
Ullman, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.
Van Boxtel, J. J., & Koch, C. (2012). Visual rivalry without spatial conflict. *Psychological Science*, 23(4), 410–418.
Van Santen, J. P. H., & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A*, *2*(2), 300–321.
Wallace, J., & Scott-Samuel, N. (2007). Spatial versus temporal grouping in a modified Ternus display. *Vision Research*, *47*(17), 2353–2366.
Werkhoven, P., Sperling, G., & Chubb, C. (1993). The dimensionality of texture-defined motion: A single channel theory. *Vision Research*, *33*(4), 463–485.
Yantis, S. (1995). Perceived continuity of occluded visual objects. *Psychological Science*, *6*(3), 182–186.
Yu, K. (2000). Can semantic knowledge influence motion correspondence? *Perception*, *29*(6), 693–707.

Chapter 100
Two Sinusoids
6 – 1 Perceptions

Jan Kremláček

PRESENTATION AND DESCRIPTION OF THE EFFECT

Illusion Construction

An interference of a stationary and moving sinusoids create ambiguous stimulus that, when observed, can result in five alternating perceptions and one that is physically plausible but hardly perceptible. Although the stimulus construction is quite simple, the perceived illusion is surprisingly strong. Since the presented stimulus is dynamic, using static book illustrations is not the best way to induce the right impression; therefore the reader should view Interactive Figure VI.100-1 on the Companion Website.

The illusion can be easily produced by a multiplication of high-frequency sinusoid (Fig. VI.100-1I) by moving low-frequency sinusoids (Fig. VI.100-1II), which can be expressed mathematically as

$$\text{ambiguous curve } (x,t) = A*\sin(f_1*x)*\ \sin(f_2*x + v*t)$$

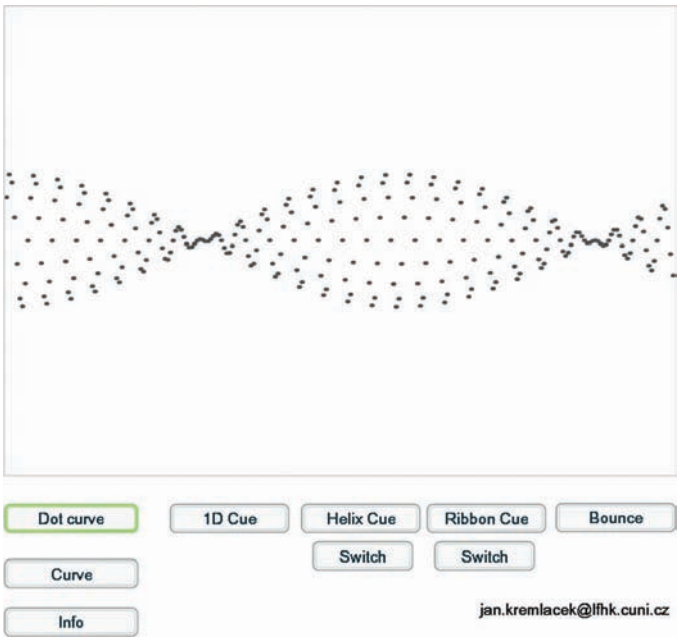
In the equation, x represents horizontal (vertical) spatial coordination, t time, f_1 higher frequency of stationary sinusoid, f_2 lower frequency of moving sinusoid, A amplitude, and v speed. Output of the equation is vertical (horizontal) spatial coordination.

Perceptions Description

Four phases of the calculated curve are depicted in Figure VI.100-1III to illustrate motion from right to left. When an observer is watching the moving pattern, up to six perceptions might emerge. Their appearance is demonstrated on a shape drawn in Figure VI.100-2I, which corresponds to *t3* phase of the curve drawn in the Figure VI.100-1III. The following perceptions could be observed:

1. Leftward-moving spindles. This is usually the first appearance an observer perceives and can be induced by fixating above or below the moving pattern or by following the pattern from right to left. This percept can be described as a flat, two-dimensional structure as depicted in Figure VI.100-2.IIa.
2. Rotating rigid double-helix. This perception might emerge when an observer moves the gaze to the right edge of the image and for a few seconds tries to follow the dots moving up and down. A three-dimensional object should emerge, as depicted in Figure VI.100-2IIb.
3. Rigid double-helix rotating in the opposite direction. Simple prolonged observing of the rotating double helix usually results in unintended reversal of the sense of the rotation.
4. Flexible ribbon with weaving edges. To observe this percept it is necessary to imagine that the pattern is a three-dimensional object, which is not the rigid rotating shape but is flexible with its edges weaving up and down like a diver's flippers. Such a shape is shown in Figure VI.100-2IIc. To facilitate this appearance in the animation, the observer might need to display external cues that emphasize the edges of the ribbon and show fixation points at the right side by pressing the "Ribbon Cue" button.
5. Overturned flexible ribbon with weaving edges. Like the rotating double-helix, the ribbon does not contain spatial cues that could uniquely determine its spatial position. Therefore it is possible to perceive that the ribbon's distant edge becomes the proximal one. To facilitate emergence of such perception, it is possible to exchange the cue emphasizing edges in the animation by pressing the "Switch" button.

   After the observer experiences the mentioned perceptions, it is possible to overturn the ribbon without an additional cue just by inducing perception of a rotating double helix; when the shape turns any odd multiple of 180° then he or she can stabilize the ribbon percept again.
6. Bouncing curves. This interpretation is hard or even impossible to perceive. The curves can be represented by dots and all aforementioned perceptions can be observed even easier. Physically the dots might be perceived as bouncing up and down without any change in lateral direction or depth. To achieve such a percept is, however, difficult because of neighboring dots bind perceptualy together along the Gestalt laws. The best way to experience this interpretation



Interactive Figure VI.100-1. Illusion demonstration. The animation will display moving dots with a simple user interface for the demonstration control. Various cues might be added to the moving structure to facilitate percepts described above in the "Perception description" part.

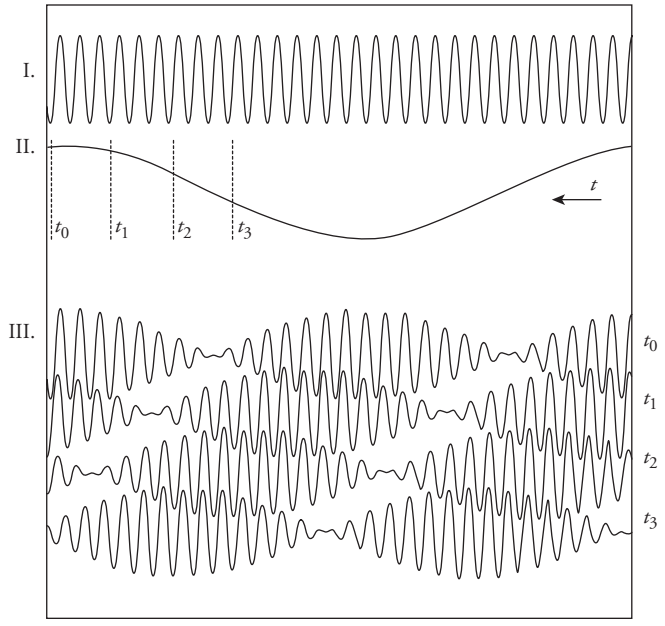


Figure VI.100-1. Illusion construction. The illusion can be easily produced just by a multiplication of high-frequency sinusoid (I) by moving low-frequency sinusoid (II). Four phases of the calculated curve are depicted (III) to illustrate motion from right to left. When an observer is watching the moving pattern, up to six perceptions might emerge. To experience the illusion, see Interactive Figure VI.100-1.

is to break the context of the scene by a mask, as is possible by pressing the "Bounce" button.

As soon as the perceptions are developed, the observer can switch among them voluntarily, or the perceptions can switch implicitly without an intention on the observer's part.

WHY THE EFFECT IS INTERESTING AND RELEVANT FOR VISION RESEARCH

For vision research, it is interesting to identify important features the brain uses to build a perception. Changing illusion parameters or adding a cue and measuring the effect on promoting some appearances offers the possibility to deduce how the visual system works and is organized.

The most astonishing part of this illusion is that using a single physical stimulus demonstrates the brain's active role in acquiring and processing visual information from two-dimensional motion to construct several possible perceptions even in three dimensions. To build a visual interpretation the brain exploits several mechanisms. In the illusion presented here, the most noticeable strategies are a kinetic depth perception and a perceptual rivalry.

Kinetic Depth Perception

Usually we perceive the spatial depth due to binocular vision. Misalignment between the left and right eyes' retinal images of an object allows our brain to compute some of the object's spatial properties. Kinetic depth perception is the ability to build the three-dimensional experience even during a monocular observation. The object or an observer has to rotate to project object's elements onto the retina with different angular speed, and then the brain, instead of spatially displaced images, integrates temporally shifted object representations to compute the object's spatial appearance. This effect had been described as an ability to see rigid three-dimensional objects instead of distorted shapes (Wallach, Connelu, & O'Connell, 1953).

Perceptual Rivalry

The other remarkable experience during observation of the presented illusion is perceptual rivalry. A famous demonstration of this phenomenon was described by Necker (1832) through a static line drawing of a cube, which can appear to an observer as viewed from above or below. Individual interpretations are not always automatically available, and learning might be necessary. Once the observer acquires a way of viewing the object, usually he or she is able to learn how to alter the perception using appropriate eye movements or gaze fixation. In the presented illusion most people easily see the rotating double-helix but cannot perceive the flexible ribbon shape. However, using cues or having someone explain how to look at the illusion can help them perceive the ribbon as well. Alterations in perception may also occur spontaneously and without any voluntary effort. Contemporary experience suggests that just one percept at a time can exist in our consciousness. This phenomenon is especially interesting for the study of visual awareness.

VARIANTS AND RELEVANT PARAMETERS FOR THE EFFECT

The presented illusion can be well perceived using only dots or two slowly moving sinusoids with the same direction and

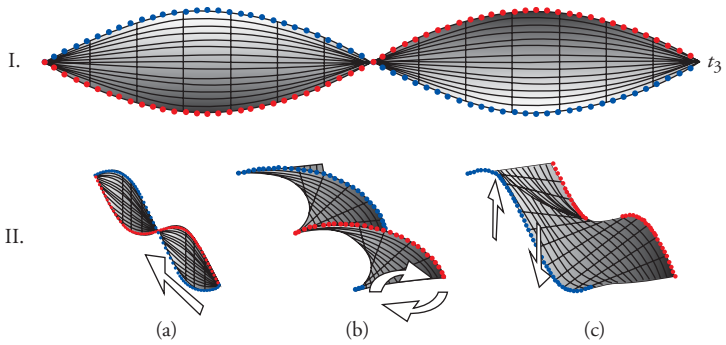


Figure VI.100-2. Three-dimensional interpretations. The shape (I) represents curve in Figure VI.100-1III. This shape allows for visualization of patterns perceived during observation of the illusion. II: (a) leftward motion of two-dimensional olives like structure; (b) rotating three-dimensional rigid double-helix; (c) three-dimensional flexible ribbon with weaving edges.

opposite amplitude. Extrinsic factors such as speed and amplitude mainly of the low-frequency sinusoid play an important role for the illusion production. When we change these factors away from the optimum, some appearances become more difficult to perceive. For example, increasing the speed v, the percept is described as leftward-moving spindles, but when we occlude the animation by an aperture and show less than a quarter of a low-frequency sinusoid, then perception of leftward motion is eliminated. Also, increasing amplitude A makes it more difficult to observe three-dimensional appearances.

Moving Sinusoids

The sinusoidal motion is so astonishing that it has appeared in many, mostly independent, observations. The first observation of a single rigid translating sinusoid transformation into a "plastic undulation of the line" was described by Nakayama and Silverman (1988). The authors found that horizontal motion of a sinusoidal curve is perceived as nonrigid if the angle at zero crossing is lower than 15°; however, this does not fully hold for the presented illusion. Simultaneous display of three translating sinusoids opposite to the barber-pole illusion as well as four perceptions during observation of single moving sinusoids were reported by Ito (1993). Similar observations of moving diamonds or rotating double helix have also been described (Gilbert, 1991).

Illusion in Other Sensory Domains

The perceptual instability of an ambiguous sensory input was observed and demonstrated in various examples. One well-known demonstration is Rubin's (1915) face-vase illusion, in which either a face or a vase is perceived. This phenomenon is not restricted only to the visual domain, and parallel auditory and tactile observations have been published. For example, when a listener tries to understand a rapidly repeated word like "rest" then he or she can also parse the auditory stream into the words "stress" or "tress" (Warren & Gregory, 1958). Presenting ambiguous tactile stimuli similar to apparent motion also produces perceptual rivalry in the somatosensory domain (Carter, Konkle, Wang, Hayward, & Moore, 2008).

DISCUSSION OF RESEARCH AND THEORIES

Brain and Kinetic Depth Perception

From anatomical and functional studies of the brain it is widely accepted that visual information is processed differently by the dorsal and ventral streams (Ungerleider & Mishkin, 1982). The dorsal stream processes more dynamic properties, like different kinds of motion and spatial information, and the ventral stream deals with object-related features such as color or shape. For general information see Kandel, Schwartz, and Jessell (2000).

In spite of a strict anatomical division of major streams' inputs on retina and subcortical levels (lateral geniculate nuclei), both streams are heavily interconnected. The aforementioned kinetic depth perception explicitly demonstrates how both streams cooperate when the analysis of visual motion is used to build a three dimensional shape. Studies of neural and metabolic responses to different moving patterns have shown that the kinetic depth perception excites, besides motion-specific areas of the dorsal stream, neural structures of the ventral stream like dorso-medial area and parieto-occipital junction typically active during perception of static objects (Paradis et al., 2000, Iwaki, Bonmassar & Belliveau, 2013).

Brain and Perceptual Rivalry

Multistable illusions present an interesting phenomenon for study of the brain, because the physical properties of stimulus are stable while perceptual interpretation is changing. This offers the option of separating neural activity related to a conscious level of processing from sensory, unconscious processing and revealing how visual awareness is built.

The alterations among perceptions can be influenced by several factors, such as eye movements, gaze fixation, and the subject's intention. Formerly it was assumed that antagonistic representations in the visual cortex compete with each other. Current opinion underlines the role of communication of higher cortical processing in prefrontal and parietal areas with visual sensory areas when building a specific percept (Leopold & Logothetis, 1999). More specifically, one can imagine that during eye movements certain features of the image on the retina stabilize in agreement with a perceptual prediction. The eye movements can

be directed by our volition or can be spontaneous. Such a view is consistent with the theory of minimization of a free energy between top-down and bottom-up sensory perceptual inference (Friston, Kilner, & Harrison, 2005).

Some studies have shown that changes in neural activity in early sensory areas (functional magnetic resonance imaging or magnetoencephalography/electroencephalography) can predict the perceptual switch; for a review see Sterzer, Kleinschmidt, and Rees (2009). Also, anatomical properties of the parietal cortex and its interindividual variability can explain observed differences in the frequency of perception shifts (Kanai, Bahrami, & Rees, 2010).

REFERENCES

Carter, O., Konkle, T., Wang, Q., Hayward, V., & Moore, C. (2008). Tactile rivalry demonstrated with an ambiguous apparent-motion quartet. *Current Biology*, *18*(14), 1050–1054.

Friston, K., Kilner, J., & Harrison, L. (2005). A free energy principle for the brain. *Journal of Physiology*, *100*(1–3), 70–87.

Gilbert, S. (1991). A new kinetic depth illusion for introductory psychology and sensation and perception courses. *Teaching of Psychology*, *18*(1), 55–56.

Ito, H. (1993). The perception of 3-D rotation from translating sine-wave lines: The reverse of the barber-pole illusion. *Perception*, *22*(2), 209–214.

Iwaki, S., Bonmassar, G., & Belliveau, J. W. (2013). Dynamic cortical activity during the perception of three-dimensional object shape from two-dimensional random-dot motion. *Journal of Integrative Neuroscience*, *12*(3), 355–67.

Kanai, R., Bahrami, B., & Rees, G. (2010). Human parietal cortex structure predicts individual differences in perceptual rivalry. *Current Biology*, *20*(18), 1626–1630.

Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2000). *Principles of neural science* (4th ed.). New York: McGraw-Hill.

Leopold, D., & Logothetis, N. (1999). Multistable phenomena: Changing views in perception. *Trends in Cognitive Sciences*, *3*(7), 254–264.

Nakayama, K., & Silverman, G. H. H. (1988). The aperture problem: I. Perception of nonrigidity and motion direction in translating sinusoidal lines. *Vision Research*, *28*(6), 739–746.

Necker, L. A. (1832). LXI. Observations on some remarkable optical phænomena seen in Switzerland; and on an optical phænomenon which occurs on viewing a figure of a crystal or geometrical solid. *Philosophical Magazine, Series 3, 1*(5), 329–337.

Paradis, A. L., Cornilleau-Peres, V., Droulez, J., Van De Moortele, P. F., Lobel, E., Berthoz, A., … Poline, J. B. (2000). Visual perception of motion and 3-D structure from motion: an fMRI study. *Cerebral Cortex*, *10*(8), 772–783.

Rubin, E. (1915). *Synsoplevede figurer*. Copenhagen: Gyldendal.

Sterzer, P., Kleinschmidt, A., & Rees, G. (2009). The neural bases of multistable perception. *Trends in Cognitive Sciences*, *13*(7), 310–318.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 586–594). Cambridge, MA: MIT Press.

Wallach, H., Connelu, D. N. O., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, *45*(4), 205–217.

Warren, R. M., & Gregory, R. L. (1958). An auditory analogue of the visual reversible figure. *The American Journal of Psychology*, *71*(3), 612.

Chapter 101
The Illusions of Numerosity

Riccardo Luccio

THE EFFECTS

The problem of the perception of the numerosity is an erratic phenomenon that originates with the beginnings of the discipline but has experienced long periods of silence. Many authors have researched it, but few have provided more than a single contribution on the subject (for a review, see Luccio, 1983). In this field, as we will see in this chapter, many phenomena have been studied, such as the effect of the relationship between the number of the elements and the area that they occupy, the regularity of the disposition of the elements, the influence of the color or the shape of the elements, the formation of subgroups, and so on. This chapter focuses on a small subfield: the so-called illusions of numerosity.

If visual illusions are mistaken perceptions that systematically induce deceptive or misleading judgments about some elements of the visually perceived images, illusions of numerosity are mistaken perceptions that induce an erroneous appreciation of the number of the elements (e.g., dots) that compose some part of the visually perceived image. We can classify the illusions of numerosity into at least two very general categories.

The first category includes illusions in which the wrong appreciation of the numerosity is accompanied by errors related to the dimensions of the percept. We have known since the beginning of the 20th century (Messenger, 1903) that with classical length illusions, such as the Müller-Lyer illusion (Fig. VI.101-1a), if we replace the lines with rows of dots, we also observe a distortion of the judgment of numerosity in the same direction of the illusion of length (see Luccio & Rodari, 1984). However, almost all illusions of length can be prepared in such a way that illusions of numerosity are also present. This is the case with the Sander illusion, the Ponzo illusion, and the vertical-horizontal illusion (Figs. VI.101-1b, VI.101-1c, and VI.101-1d, respectively), but there are lesser-known illusions, such as Botti's (1909) illusion (Fig. VI.101-1f), that exhibit very impressive effects. The same is true with the illusions of area, such as Delbœuf's illusion (Fig. VI.101-1e).

Notice that the correlation between the overestimation of numerosity and the overestimation of size (length, area) is not always present. Alam, Luccio, and Vardabasso (1986) and Agostini and Luccio (1994) observed that the judgment of numerosity depends also on the method used: in simultaneous comparison, we observe a relative overestimation compared with direct estimation. The paradox noticed by Agostini and Luccio results from the fact that with the method of constant stimuli, both length and numerosity judgments move in the same direction, whereas using a magnitude-estimation method, the two illusions move in opposite direction. This fact was observed many years ago by De Marchi (1929), who, however, did not provide experimental data that shows that when the participants are asked to perform a constant stimuli estimation versus a magnitude estimation task, the illusion of numerosity moves in opposite directions. Notice that constant stimuli method involves a relative judgment, with a comparison between patterns, while magnitude estimation involves an absolute judgment of each pattern per se. In more recent research, I have confirmed these previously paradoxical results with other patterns (regular and irregular patterns, Botti's illusion, vertical-horizontal illusion; see Luccio, 2011).

The second category includes illusions in which the error of appreciation of the numerousness does not accompany any remarkable perceptual distortion. It deals with elements arranged in two patterns and in which the number of the elements of a pattern is overestimated in comparison to that of the other pattern. Sometimes the elements are in direct one-to-one correspondence from a pattern to the other, but this does not help to appreciate its equal numerosity. For example compare the illusion of Ponzo (1928; Fig. VI.101-2a), with that of Pegrassi (1904; Fig. VI.101-2b), with that of the "solitary" of Frith and Frith (1972; Fig. VI.101-2c). The Ponzo illusion is particularly interesting: the elements consist of four in each pattern, yet the viewer observes: "The number is equal, but the elements of the superior row are more." It is clear that one most invoke different mechanism to explain these two categories of illusions.

THE RELEVANCE OF THE ILLUSIONS

In my opinion, studies on the perception of the numerosity have been neglected for many years in error. Notice that these studies were initiated particularly with the research on the span of apprehension, or *Umfang* (Cattell, 1885), which were completely forgotten up to World War II. It

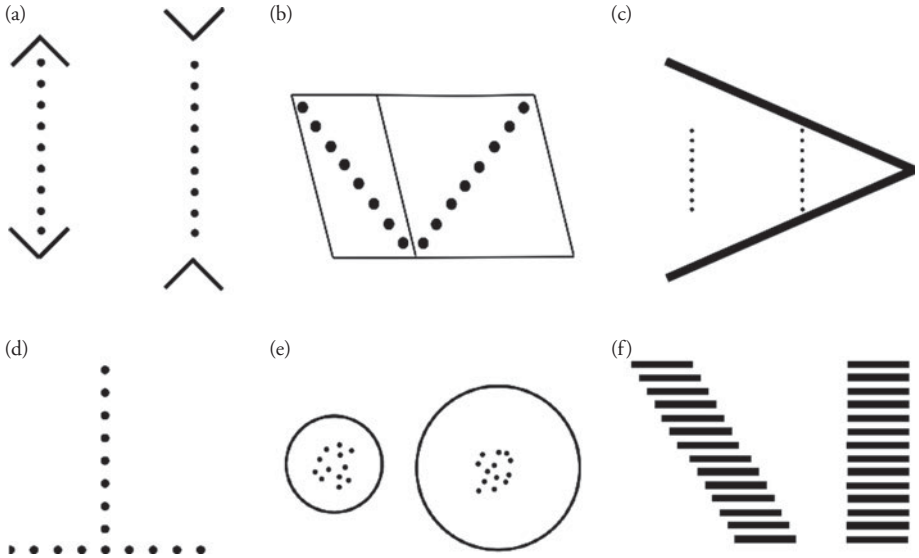


Figure VI.101-1. Examples of illusions of length or area transformed in illusions of numerosity. (a) Variant of the Müller-Lyer illusion. The dots on the row on right appear more numerous. (b) Variant of the Sander illusion. The dots on the diagonal on right appear more numerous. (c) Variant of the Ponzo illusion. The dots on the row on right appear more numerous. (d) Variant of the vertical-horizontal illusion. The dots on the vertical row appear more numerous. (e) Variant of Delbœuf's illusion. The dots in the circle on right appear less numerous. (f) Botti's illusion. The lines on the left appear more numerous.

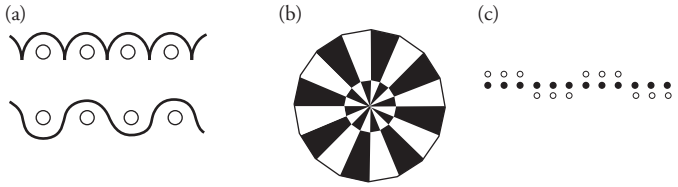


Figure VI.101-2. (a) The Ponzo illusion. The dots on the top have clearly the same number as the dots on the bottom; nevertheless there appear to be more. (b) Pegrassi's illusions. The elements in the center of the figure appear more numerous, despite the correspondence one-to-one with the external ones. (c) Variant of the solitaire illusion. The white dots appear less numerous, despite the correspondence one-to-one with the black ones (modifications of the original ones by Frith & Frith).

was then that concepts such as subitizing, estimating, and counting were established with Kauffmann, Lord, Reese, and Volkman (1949), which have been of conclusive importance for cognitive psychology during the past half-century.

I believe that the relative neglect of the illusions of numerousness constitutes, however, of an error of the identical type. It is well known that, for all the illusions, the perceptual mechanisms in action are exactly the same ones that come into play in daily perception, but the illusions are precious because they show the actions of these mechanisms in extreme situations. In our case, of course, there are some specific problems relative to the cognitive mechanisms underlying the concept of numerosity.

The first problem is the relationship between the concept of numerousness and numbers. The dissociation between direct comparison and estimation that we observe in the illusions of type 1, as well as the very results of the illusions of type 2, demonstrate clearly that one can assign a certain number to a collectivity of elements, but that number does not represent necessarily the "numerousness" of the collectivity. This fact was demonstrated many years ago by Piaget and Szeminska (1941): if we present a child with two rows containing the same numbers of tokens but in one row the tokens are more spaced, so the row is longer, the child will report that the number in the rows is the same but in the longer row there are more tokens. Similar behaviors were demonstrated in "primitive" peoples (*Naturvölker*) by Wertheimer (1912), exactly one century ago, at the very beginning of Gestalt psychology.

The second relevant problem concerns the relationship between the space occupied by the elements that one must judge and the perceived numerousness. Here the data are contradictory. In general, the elements appear to be more numerous when the space that they occupy is larger, but this is not always the case (see Figs. VI.101-2b and VI.101-2c).

Much data have been collected about the strategies that participants apply in assessing numerousness judgments. In this respect, however, there is a total lack of research in the field of the illusions. Regardless, there is no reason to suspect that the strategies that participants use in this field are any different from those used in other fields.

VARIANTS AND PARAMETERS

In discussing the possible variants of the illusory patterns, it could be useful to distinguish between two different kinds of parameters: order parameters and control parameters. This distinction, customary in other fields of knowledge (for example, in the discussion of the dynamic systems; see Haken, 1990), refers respectively to the global and the local properties of the patterns. The order parameter is a quantity that has value zero in presence of a certain configuration, but if there is a phase transition, assumes a non-zero value. For example, in the Müller-Lyer illusion if we invert the direction of the arrows from inward to outward (control

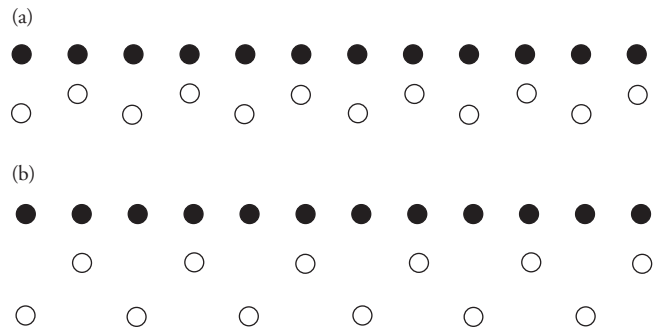


Figure VI.101-3. Variants of the solitaire illusion. In (a) the white dots appear more numerous; in (b) they appear less numerous. (Modifications of the original ones by Frith & Frith.)

parameter) we have a phase transition, and what was underestimated now is overestimated; and vice versa.

Of course we could be interested in the control parameters, and in the limits in which they can change while leaving the order parameter unchanged, and how this can affect the illusion of numerosity. From this point of view, however, there is an almost complete absence of empirical evidence. We have more data about the changing of the order parameters (i.e., a phase transition) consisting of an inversion of the illusion. Particularly interesting here is the case of the variants of the solitaire illusion (Fig. VI.101-2c).

In Figure VI.101-3 we have two variants of the illusion proposed by Luccio (1981). In the original figure observers underestimate the number of white dots, half of which are on one side of the row of black dots and half of which are on the other side. In Figure VI.101-3a, all the white dots are in two rows on the same side of the row of the black dots: here the illusion is clearly inverted, and the white dots are relatively overestimated. Now, however, if we move away the two rows of white dots, we have another phase transition, and the white dots appear to be fewer than the black ones (Fig. VI.101-3b). Of course the area occupied by the white dots is larger in Figure VI.101-3a than in Figure VI.101-3b, and this is in apparent contradiction with the well-known fact that the greater the occupied area, the more numerous the elements that occupy it appear (Allik & Tuulmets, 1991). We discuss this fact later.

Other control parameters of interest could be (a) the exposure time, (b) the form of the rows of the elements, (c) the overall size of the patterns, and (d) the relative size of the parts of the pattern. The importance of these possible variants is discussed later, but one must stress that, until now, we have had no evidence at all about them in regard to illusions.

DISCUSSION

As discussed earlier, the first problem to consider is the discrepancy between the judgments given with psychophysical methods that involve a direct comparison between a standard and a variable (as with the method of constant stimuli, the staircase method, and so forth) and direct measurements, as in magnitude estimation. In the latter case, the judgment is always closer to the true number of elements, and the illusory effect tends to disappear; in some cases, the effect is inverted. According to Luccio (1983), these results can be explained if we consider that two independent processes act at the basis of the judgment of numerosity: a process of *estimating*, which acts in direct measurement when the perceptual input goes through a quantitative internal representation, and the relative *impression of power* (*Mächtigkeitseindruck*; Liebenberg, 1914), which acts in comparing two patterns, before the formation of the internal quantitative representation, which appears unnecessary in such a condition. In other words, from one side there is a relatively complex process that requires a variety of complex strategies (see later discussion) and needs some time to be accomplished. From the other side, the impression of power is a physiognomic (tertiary) property of the percept that is immediately caught by the observer, before any attempt to give a determination of the exact number of the elements. As Metzger (1954) said, it is a property that has the quality of the *met* (*Angetroffene*).

Of course I am aware of the difficulty of giving an operational definition of *Mächtigkeitseindruck*. We can say that this impression is greater (and the correlative judgment of numerosity is greater) when (a) *ceteris paribus*, the elements occupy a larger space; (b) the elements appear more dense; (c) the elements are greater; (d) the shape of the elements is complex and more articulated; and so on. In the illusions we are analyzing here, factors (a) and (b) appear to be the most important.

Usually, in the illusions of type 1 (Fig. VI.101-1), factor (a) prevails (i.e., the elements are put on longer rows or in larger spaces, or at least in spaces that are perceived as longer or larger). In the illusions of type 2 (Fig. VI.101-2), factor (b) prevails (i.e., the elements appear more dense).

Factors (a) and (b) can be put into conflict. It is clear, from Pegrassi's illusion (Fig. VI.101-2b), that the elements in the center of the figure appear more numerous because are more dense, although the area they occupy is smaller. Notice that in the Ponzo illusion (Fig. VI.101-2a) density is again the factor at play, but here the greater density is given by irrelevant elements in the pattern.

In my opinion, it makes little sense to consider the size of the area occupied by the elements to be estimated a factor of overestimation, and in this sense it could be misleading to speak about "numerosity illusion" with reference to the perceptual interaction between elements and their spatial arrangement, or to the perceptual clustering of elements based on their relative proximity, without further specifications (cf. Vos, Van Oeffelen, Tibosch, & Allik, 1988). Too many factors act that are separately controlled: the connectedness of the elements (e.g., He, Zhang, Zhou, & Chen, 2009), their segmentation (e.g., Franconeri, Bemis, & Alvarez, 2009), the regularity versus randomness of their disposition (Ginsburg, 1978); the shape of the rows on which they are placed (Krishna & Raghubir, 1997); and so on, A particularly interesting field of research that has opened up in the past few years involves the spatial representations of numbers and numerical concepts (see de Hevia & Spelke, 2009; Hubbard, Piazza, Pinel, & Dehaene, 2005). All of these factors, and many others that cannot be enumerated here, are as yet far from constituting an integrated framework in which the problem of numerosity illusions is also posited. They not only influence the impression of power given by the pattern but, when the

task is estimating, determine which strategies the participants employ. As well documented by Luwel, Verschaffel, Onghena, and De Corte (2003a, 2003b), the participants use several strategies to determine the number of the elements of a pattern and are highly flexible when adapting them to the task's demands. However, in this case we also lack a comprehensive framework.

In conclusion, the field of the illusions of numerosity is a neglected one, and we regret this, because here we can observe many beautiful effects that have not been explained and that deserve more extensive research. What is particularly regrettable is that we lack an integration between this field and the larger one of the perception of numerosity.

REFERENCES

Agostini, T., & Luccio, R. (1994). Müller-Lyer illusion and perception of numerosity. *Perceptual and Motor Skills*, *78*, 937–938.

Alam, S., Luccio, R., & Vardabasso, F. (1986). Regularity, exposure time, and perception of numerosity. *Perceptual and Motor Skills*, *63*, 883–888.

Allik, J., & Tuulmets, T. (1991). Occupancy model of perceived numerosity. *Perception & Psychophysics*, *49*, 303–314.

Cattell, J. M. (1885). Über die Zeit und Erkennung und Benemung von Schriftzeichen, Bildern und Farben. *Wundts Philosophischen Studien*, *2*, 635.

De Marchi, S. (1929). Illusioni ottiche e apprezzamento di collettività. In *Atti IX Congresso Psicologi Italiani.* (pp. 8–18), Padova, Cedam.

de Hevia, M. D., & Spelke, E. S. (2009). Spontaneous mapping of number and space in adults and young children. *Cognition*, *110*(2), 198–207.

Franconeri, S. L., Bemis, D. K., & Alvarez, G. A. (2009). Number estimation relies on a set of segmented objects. *Cognition*, *113*, 1–13.

Ginsburg, N. (1978). Perceived numerosity, item arrangement, and expectancy. *The American Journal of Psychology*, *91*, 267–273.

Haken, H. (1990). Synergetics as a tool for the conceptualisation and mathematisation of cognition and behaviour: How far can we go? In H. Haken & M. Stadler (Eds.), *Synergetics of cognition* (pp. 2–30). Berlin: Springer.

He, L., Zhang, J., Zhou, T., & Chen, L. (2009). Connectedness affects dot numerosity judgment: Implications for configural processing. *Psychonomic Bulletin & Review*, *16*, 509–517.

Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews and Neuroscience*, *6*, 435–448.

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkman, J. (1949). The discrimination of visual number. *The American Journal of Psychology*, *62*, 498–525.

Krishna, A., & Raghubir, P. (1997). The effect of line configuration on perceived numerosity of dotted lines. *Memory & Cognition*, *25*, 492–507.

Liebenberg, R. (1914). Über das Schätzen von Mengen. *Zeischrift für Psychologie*, *68*, 321–395.

Luccio, R. (1981). Formazioni di sottogruppi e percezione di numerositü. *Archivio di Psicologia, Neurologia e Psichiatria*, *42*, 463–476.

Luccio, R. (1983). La percezione della numerosità. *Storia e Critica della Psicologia*, *4*, 215–283.

Luccio, R. (2011). Direct estimation and simultaneous comparison in perception of numerosity: A paradox. In *Zoran & Ramiro Bujas days*. Zagreb.

Luccio, R., & Rodani, M. G. (1985). Illusioni di lunghezza e illusioni di numerosità. In G. Kanizsa (Ed.), *Fenomenologia sperimentale della visione* (pp. 259–265). Milano: Angeli.

Luwel, K., Verschaffel, L., Onghena, P., & De Corte, E. (2003a). Analysing the adaptiveness of strategy choices using the choice-no-choice method. *European Journal of Cognitive Psychology*, *15*, 511–537.

Luwel, K., Verschaffel, L., Onghena, P., & De Corte, E. (2003b). Flexibility in strategy use: Adaptation of numerosity judgment strategies to task characteristics. *European Journal of Cognitive Psychology*, *15*, 247–266.

Messenger. (1903). The perception of number. *Psychological Review Monographs*, *5*, n. 5.

Metzger, W. (1954). *Psychologie.* Darmstadt, Germany: Steinkopf.

Piaget, J., & Szeminska, A. (1941). *La génèse du nombre chez l'enfant Neuchâtel*. Paris: Delachaux & Niestlé.

Ponzo, M. (1928). Le illusioni negli apprezzamenti di collettività. *Archivio Italiano di Psicologia*, *6*, 1–37.

Vos, P. G., Van Oeffelen, M. P., Tibosch, H. J., & Allik, J. (1988). Interactions between area and numerosity. *Psychological Research*, *50*, 148–154.

Wertheimer, M. (1912). Über das Denken der Natuvölker. *Zeitschrift für Psychologie*, *60*, 321–368.

Chapter 102
The Aperture Capture Illusion

Evan M. Palmer and Philip J. Kellman

INTRODUCTION

Compared to earlier eras, contemporary work in visual perception includes a greater recognition of the importance of events, or changing information in the environment, particularly during object and observer motion. These functional priorities of perception depend on information extended in time, including relations not available in any static image, since dynamic information is often fragmentary or incomplete in any given moment.

Researchers in recent decades have made great progress in understanding the visual processes that allow perception of the persisting objects, surfaces, layouts, and events from information in motion and change. As with other areas of perception, accompanying a greater understanding of basic processes and functions are new and significant visual illusions. Here we describe the aperture capture illusion, a vivid, counterintuitive set of effects that turns out to have deeply revealing connections to processes we use to perceive coherent objects from information that is fragmentary in both space and time (see Palmer & Kellman, 2014).

PRESENTATION AND DESCRIPTION OF THE EFFECT

When a vertical rod passes behind an occluding surface with two misaligned apertures, the whole, connected, aligned rod appears broken, disconnected, and misaligned (Fig. VI.102-1 and Video VI.102-1).

The apparent misalignment of the rod pieces is determined by the arrangement of the apertures rather than the motion direction of the rod (Videos VI.102-2 and VI.102-3). Regardless of whether the rod moves leftward or rightward, the pieces appear to be misaligned in the same direction as the misalignment of the apertures. Hence the name "aperture capture illusion."

The apparent misalignment of the rod's pieces can be counteracted by shifting one piece of the rod to null the illusion. Thus a misaligned rod may appear aligned as it passes behind an occluding surface and is seen through two misaligned apertures (Video VI.102-4). The magnitude of misalignment needed to null the illusion can be recorded as a dependent variable, allowing the strength of the illusion under various circumstances to be determined.

Higher velocities of translation result in greater apparent misalignment of the rod (Videos VI.102-5, VI.102-6, and VI.102-7), necessitating larger magnitudes of compensatory shifting to null the illusion. The magnitude of the illusion increases monotonically with velocity (Fig. VI.102-2).

Interestingly, the illusion occurs even when the rod is seen through two apertures simultaneously, as long as one aperture is longer than the other (Video VI.102-8). The subjective impression one has in Video VI.102-8 is that the rod appears aligned for a moment while both the top and bottom pieces are visible and then becomes misaligned as the top piece of the rod is occluded before the bottom piece. This observation suggests two important components of the illusion.

First, moving objects continue to be visible for a short time after they become occluded, allowing for perceptual integration of visible and occluded regions (Palmer, Kellman, & Shipley, 2006; see later discussion). Second, the perceived velocity of occluded regions of an object is slower than their true velocity. If both assertions are true, then an essential element of the illusion is that one portion of the rod becomes occluded before the other part. The occluded region remains part of the observer's perceptual experience but appears to be moving more slowly than the visible region, resulting in an apparent misalignment. This theory also explains why the illusion can be nulled. Under this explanation, a prediction is that if both visible regions of the rod become occluded at the same time, then one should see no illusion. This prediction is confirmed in Video VI.102-9, which merely reverses the arrangement of apertures from Video VI.102-8.

WHY THE EFFECT IS INTERESTING AND RELEVANT FOR VISION RESEARCH

The aperture capture illusion, like many illusions, reveals the workings of processes that produce accurate outcomes in typical circumstances but in other cases lead to a bizarre percept. In this case, the paucity of apertures through which to view the translating object, as well as the misalignment of the apertures, causes the object to appear broken and misaligned in the direction of the apertures. What perceptual processes are in play here, and under what circumstances might they operate normally?

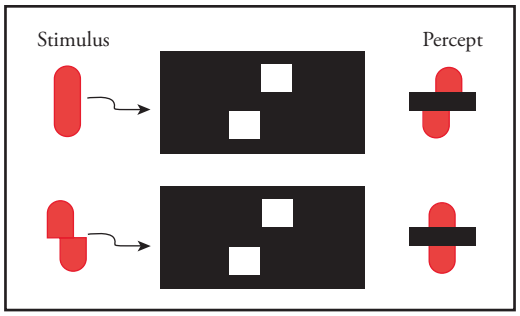
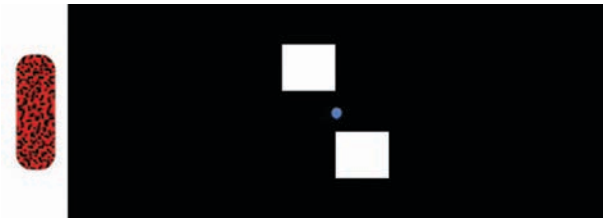
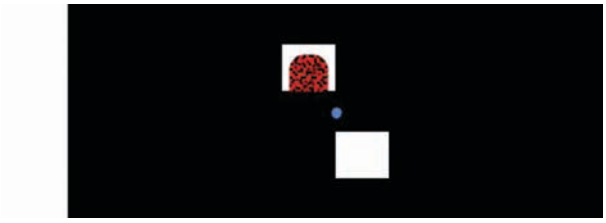


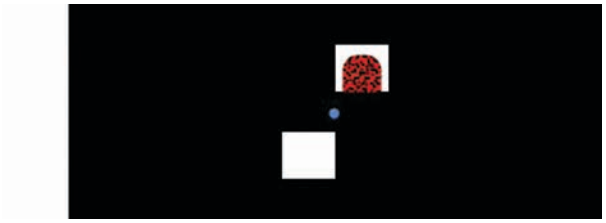
Figure VI.102-1. A schematic depiction of the aperture capture illusion. Top: When an aligned rod moves behind an occluding surface with two misaligned apertures, the rod appears to be misaligned in the same direction as the offset of the apertures. Bottom: The percept of misalignment can be nulled by misaligning the rod in the direction opposite to the offset of the apertures. (Figure by Palmer & Kellman, 2014.)



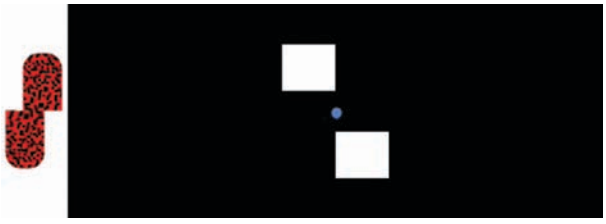
Video VI.102-1. Figure by Evan M. Palmer and Philip J. Kellman.



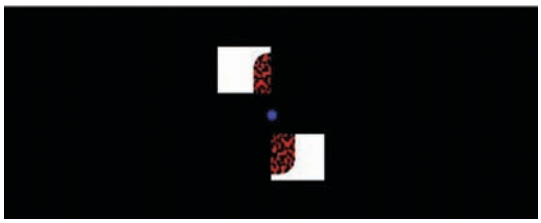
Video VI.102-2. Figure by Evan M. Palmer and Philip J. Kellman.



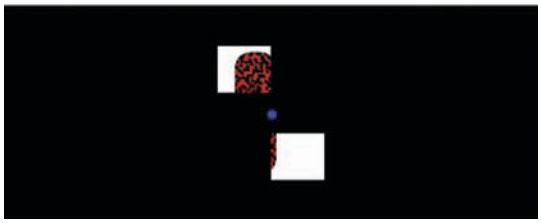
Video VI.102-3. Figure by Evan M. Palmer and Philip J. Kellman.



Video VI.102-4. Figure by Evan M. Palmer and Philip J. Kellman.



Video VI.102-5. Figure by Evan M. Palmer and Philip J. Kellman.



Video VI.102-6. Figure by Evan M. Palmer and Philip J. Kellman.



Video VI.102-7. Figure by Evan M. Palmer and Philip J. Kellman.

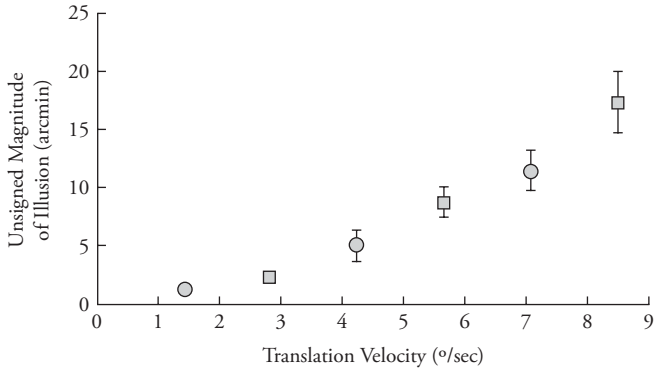
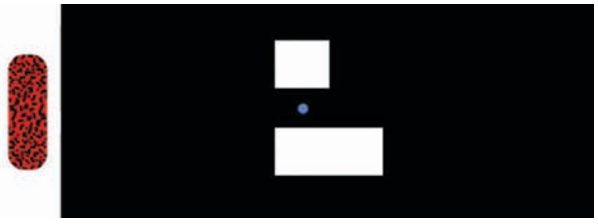


Figure VI.102-2. Magnitude of the illusion as a function of rod velocity. Data are the absolute value of the final (mis)alignment, averaged across motion direction and aperture configuration. Error bars represent ± one standard error of the mean. (Figure by Palmer & Kellman, 2014.)



Video VI.102-8. Figure by Evan M. Palmer and Philip J. Kellman.



Video VI.102-9. Figure by Evan M. Palmer and Philip J. Kellman.

Many objects in the world are occluded, and when those objects move, their visible regions can change dramatically over time. The challenge to the human visual system is to collect and integrate these dynamically occluded visible regions into a percept that supports perception and action. Under ecologically normal circumstances, dynamically occluded objects are seen through surfaces with many apertures, as when viewing a fire truck driving down the street through the branches of a tree (Video VI.102-10). In these situations, many regions of the object are visible simultaneously, and the perception of a whole, complete, spatially coherent object is robust. However, when dynamically occluded objects are seen through just two apertures, spatiotemporal shape integration breaks down, resulting in the aperture capture illusion.

To understand how the visual system perceives dynamically occluded objects, it is important to first understand how it perceives statically occluded objects. Kellman and Shipley (1991) proposed the geometric principle of *relatability* to describe the circumstances under which occluded edges will appear to connect behind an interposed surface. Under a wide range of circumstances, relatability expresses a necessary condition on contour interpolation (for both occluded and illusory contours). Interpolation occurs between surface edge fragments separated by a gap. When one object's bounding contours are partially occluded by another object, some visible edges of the occluded object will lead into tangent discontinuities (contours junctions). In illusory contour displays, there is no visible occluder, but the edge fragments must lead into a sharp corner (Shipley & Kellman, 1990).



Video VI.102-10. Figure by Evan M. Palmer and Philip J. Kellman.

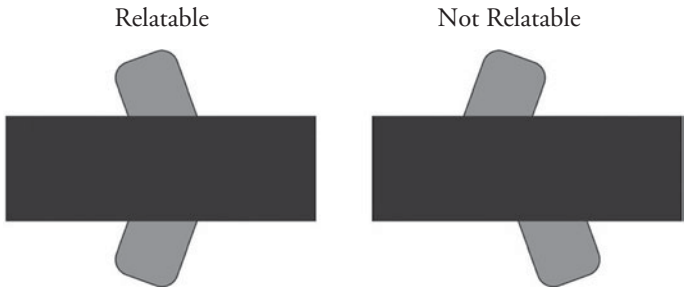


Figure VI.102-3. Examples of relatable and nonrelatable figures. The gray object on the left has relatable edges and appears to connect behind the black occluding surface, but the two pieces of the gray object on the right are not relatable and thus do not appear to connect behind the occluder. The edges of the gray object on the right are not relatable because they cannot be connected by a smooth curve that only bends once and forms an obtuse angle between the edges. (Figure by Evan M. Palmer and Philip J. Kellman.)

Relatability is a condition on the geometric relations of two surface edge fragments ending in contour junctions. Briefly, if two such edge fragments can be connected by a smooth, monotonic contour that bends through no more than 90°, the edge fragments are relatable (Fig. VI.102-3). Visible surface edges that satisfy these conditions support interpolation—connection across gaps—allowing complete objects to be perceived despite spatial gaps in the input (Kellman, Garrigan, & Shipley, 2005). Relatability as described provides a formal account of the spatial relations in static displays that lead to contour interpolation.

When objects move and are dynamically occluded, the input edges to interpolation processes may not be visible at the same time or in the same place. Instead, the edges may be revealed sequentially, in different parts of the visual field. How does the visual system unite the regions of a dynamically occluded object that twinkle into and out of view through an intervening occluded surface? Palmer et al. (2006) broadened the notion of relatability into the more comprehensive account of *spatiotemporal relatability* (STR) by adding two assumptions.

The first assumption of STR is *persistence*. When moving objects pass behind an occluding surface and are thus physically out of view, they nonetheless continue to briefly have a perceptual presence similar to the visual icon representation first described by Sperling (1960). Occluded but persisting object fragments may be perceptually integrated with visible regions, allowing the visual system to achieve visual unit formation despite temporal fragmentation of the visual input.

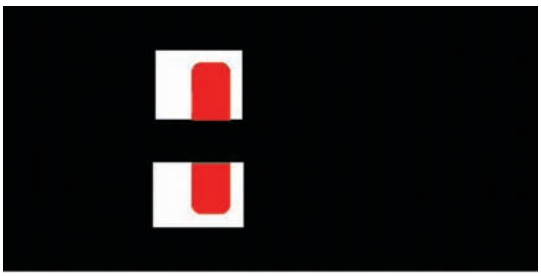
Second, the notion of *position updating* proposes that the relative positions of occluded but persisting object fragments are updated over time, in accord with their previously observed motion velocity. Thus both occluded and visible regions of a dynamically occluded object stay in spatial register with each other, creating an object-centered reference frame. Within this moving reference frame centered on the dynamically occluded object, relatability determines which edges become integrated into perceptual wholes behind the occluding surface. In its normal operation, STR allows the visual system to gather sequential information about object fragments, use their

velocity signals to project their future positions, and perform interpolation (filling-in) between relatable edges of currently visible fragments and previously stored, positionally updated fragments.

Considering the aperture capture illusion through the framework of STR leads to some interesting insights. First, the fact that a unified percept of the rod (misaligned or not) can be achieved through two spatially separated apertures (e.g., Video VI.102-1) gives support to the notions of persistence and position updating. Second, the idea that occluded regions of the object appear to move more slowly than visible regions (e.g., Video VI.102-8) suggests a flaw in the position updating process. Specifically, it appears that position updating is feasible but unreliable in the sense that the motion vector of occluded regions is underestimated (Palmer & Kellman, 2002). Under more ecological conditions, as with the fire truck in Video VI.102-10, apertures are plentiful and closely spaced, meaning that persistence and position updating do not have to operate over large spatial or temporal gaps. Regions of the dynamically occluded object do not go out of sight for long, meaning that position updating only has to operate for the briefest of moments. In the aperture capture illusion, persistence and position updating have to operate for much longer, revealing a flaw in the position updating process that is otherwise not normally apparent.

VARIANTS AND RELEVANT PARAMETERS

- Direction of object motion does not seem to matter (Video VI.102-2), but the arrangement of apertures does (compare Video VI.102-2 with Video VI.102-3; Palmer & Kellman, 2014).
- As the velocity of the object increases, the size of the illusion also increases (Videos VI.102-5, VI.102-6, and VI.102-7; Fig. VI.102-2; Palmer & Kellman, 2014).
- The illusion does not occur when the two fragments are seen through two aligned apertures, but it does occur when one aperture is longer than the other (even if the object is seen briefly through two apertures simultaneously; Video VI.102-11; Palmer & Kellman, 2014).
- Increasing the size of the first aperture in which the rod appears does not eliminate the illusion, suggesting that the illusion is not the result of improper velocity perception of the rod before occlusion (Palmer, 2003; Palmer & Kellman, 2014).



Video VI.102-11. Figure by Evan M. Palmer and Philip J. Kellman.

- As the space between the apertures increases, the size of the illusion also increases (Palmer, 2003; Palmer & Kellman, 2014).
- If the rod is stationary and the occluder moves in front of it, there is no perceived misalignment if the observer fixates on the rod, but there is if the observer optically pursues the moving occluder (Palmer, 2003).
- When there are many closely spaced apertures, the illusion does not occur (see displays in Palmer et al., 2006).

DISCUSSION OF RESEARCH AND THEORIES

The aperture capture illusion is interesting because it represents a failure of proper shape integration from dynamically occluded objects. Under normal circumstances, perception of dynamically occluded objects is robust (Keane, Lu, & Kellman, 2007; Palmer et al., 2006), as is the case in static scenes when requisite stimulus conditions are satisfied (e.g., Bregman, 1981; Kellman et al., 2005; Kellman & Shipley, 1991), but in this stimulus situation with just two apertures, perception is inaccurate (Palmer & Kellman, 2014). The processes of persistence and position updating in the theory of STR (Palmer et al., 2006) are both suggested by prior research. The notion of persistence is supported by the visual icon representation, first discovered by Sperling (1960) and named by Neisser (1967). The visual icon is a detailed and faithful representation of stimulus that is no longer visible, retaining stimulus properties such as color (Banks & Barber, 1977; Clark, 1969), stimulus motion and direction (Demkiw & Michaels, 1976; Shioiri & Cavanagh, 1992), and stimulus size and orientation (Von Wright, 1968).

The notion of position updating as proposed in the theory of STR is new, but clues to its existence are scattered throughout the literature. Hochberg (1968) performed studies in which the edges of a plus-shaped figure were visible sequentially through a circular aperture and showed that observers could discern whether the figure was a possible or an impossible object. This perceptual discrimination would be impossible unless observers could keep track of the relative alignment of the edges during the successive views of the object. Michotte (1950) and others reported classic studies on the tunnel effect, a perceptual illusion in which an object appears to continue behind an occluding surface and emerge from the other side. In reality one object moves behind the surface and another identical-looking object emerges, but if the timing between these two events is correct, the percept of a single object passing behind the occluding surface is achieved.

The aperture capture illusion (in which a moving object is seen through two large, misaligned apertures) is similar to anorthoscopic perception displays (in which a moving object is seen through a single, narrow aperture). Plateau (1836) first coined the term "anorthoscopic," which means "abnormally viewed" to describe the strange percept of seeing a whole object through a narrow slit. A more ecological version of this situation might be trying to perceive a colleague walking past one's office through the narrow opening in the door. As opposed to the aperture capture

illusion, anorthoscopically perceived objects appear larger than the slit through which they were seen (Anstis & Atkinson, 1967; Parks, 1965; Zöllner, 1862). However, similar to the aperture capture illusion, anorthoscopically viewed objects appear farther along their motion trajectory than where the aperture was located (Sohmiya & Sohmiya, 1994) and appear to be a whole unified object despite only a fraction of the object being visible at once (Parks, 1965; Rock, Halper, DiVita, & Wheeler, 1987). In the aperture capture illusion case, the object is fractionated between two apertures rather than within a single aperture.

It is important to note that the aperture capture illusion cannot be explained by the notion of "retinal painting," first described by Helmholtz (1867/1962) and sometimes advanced as an explanation for anorthoscopic perception (e.g., Anstis & Atkinson, 1967). Under this account, the observer performs a smooth pursuit eye movement that tracks an object as it passes behind intervening surfaces. As fragments of the dynamically occluded object pass into and out of view, they are "painted" onto the back of the observer's retinas, allowing perception of the complete form. First, notice that this explanation relies on the idea of visible persistence, as advanced by the theory of STR. However, under the retinal painting account, position updating is accomplished via smooth pursuit eye movements rather than some other form of visuospatial processing. Considering the version of the aperture capture illusion depicted in Video VI.102-1, if the observer smoothly pursues the moving rod with the eyes, then the retinal painting hypothesis predicts that there should be no illusion. Yet experimental data show that the illusion still occurs under pursuit conditions (Palmer & Kellman, 2002, 2014). On the other hand, if the observer fixates on a single location in the display during the animation, then the retinal painting hypothesis predicts that the illusion should *never* be able to be nulled since the two pieces of the object always appear in different places on the retina. This prediction has been disconfirmed in numerous experiments in which observers have consistently been able to eliminate the illusion by suitable spatial arrangement of the visible fragments. Thus, with either central fixation or smooth pursuit, the retinal painting hypothesis is an unsatisfactory explanation for the existence of the aperture capture illusion and the ability to null it.

Another area of research with potential connections to the aperture capture illusion is multiple object tracking through occlusion. Scholl and Pylyshyn (1999) established that observers can track multiple objects through short episodes of occlusion if the optical transformations signaling occlusion are realistic (containing accretion/deletion cues; cf. Gibson, Kaplan, Reynolds, & Wheeler, 1969) rather than abnormal (imploding/exploding at the point of occlusion). Furthermore, Flombaum, Scholl, and Pylyshyn (2008) showed that observers actually pay *more* attention to the positions of tracked objects during occlusion, suggesting a persisting perceptual representation of those objects. However, despite increased attention to occluded objects during multiple object tracking, the positions of occluded objects are not well represented (Franconeri, Pylyshyn, & Scholl, 2012; Keane & Pylyshyn, 2005; Scholl & Nevarez, 2002), consistent with the aperture capture illusion.

REFERENCES

Anstis, S. M., & Atkinson, J. (1967). Distortions in moving figures viewed through a stationary slit. *Quarterly Journal of American Psychology*, *80*, 572–585.

Banks, W. P., & Barber, G. (1977). Color information in iconic memory. *Psychological Review*, *84*, 536–546.

Bregman, A. (1981). Asking the "what for" question in auditory perception. In M Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (p. 99–118). Hillsdale, NJ: Lawrence Erlbaum.

Clark, S. E. (1969). Retrieval of color information from preperceptual memory. *Journal of Experimental Psychology*, *82*, 263–266.

Demkiw, P., & Michaels, C. F. (1976). Motion information in iconic memory. *Acta Psychologica*, *40*, 257–264.

Flombaum, J. I., Scholl, B. J., & Pylyshyn, Z. W. (2008). Attentional resources in tracking through occlusion: The high-beams effect. *Cognition*, *107*, 904–931.

Franconeri, S. L., Pylyshyn, Z. W., & Scholl, B. J. (2012). A simple proximity heuristic allows tracking of multiple objects through occlusion. *Attention, Perception & Psychophysics*, *74*(4), 691–702.

Gibson, J. J., Kaplan, G. A., Reynolds, H. N., & Wheeler, K. (1969). The change from visible to invisible: A study of optical transitions. *Perception & Psychophysics*, *5*, 113–116.

Helmholtz, H. von (1962). *Treatise on physiological optics*. Vol. III ((3d ed.). New York: Dover. (Original work published 1867)

Hochberg, J. (1968). In the mind's eye. In R. N. Haber (Ed.), *Contemporary theory and research in visual perception* (pp. 309–331). New York: Holt, Rinehart & Winston.

Keane, B. P., Lu, H., & Kellman, P. J. (2007). Classification images reveal spatiotemporal interpolation in illusory figures. *Vision Research*, *47*, 3460–3475.

Keane, B. P., & Pylyshyn, Z. (2005). Is motion extrapolation employed in multiple object tracking? Tracking as a low-level, non-predictive function. *Cognitive Psychology*, *52*(4), 346–368.

Kellman, P. J., Garrigan, P., & Shipley, T. F. (2005). Object interpolation in three dimensions. *Psychological Review*, *112*(3), 586–609.

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, *23*, 141–221.

Michotte, A. (1950). On phenomenal permanence: Facts and theories. *Acta Psychologica*, *7*, 293–322.

Neisser, U. (1967). *Cognitive psychology*. East Norwalk, CT: Appleton-Century-Crofts.

Palmer, E. M. (2003). *Spatiotemporal relatability in the perception of dynamically occluded objects* (Unpublished doctoral dissertation). University of California, Los Angeles.

Palmer, E. M., & Kellman, P. J. (2002). Underestimation of velocity after occlusion causes the aperture-capture illusion [Abstract]. *Journal of Vision*, *2*, 477a.

Palmer, E. M., & Kellman, P. J. (2014). The aperture capture illusion: Misperceived forms in dynamic occlusion displays. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(2), 502–524.

Palmer, E. M., Kellman, P. J., & Shipley, T. F. (2006). A theory of dynamic occluded and illusory object perception. *Journal of Experimental Psychology: General*, *135*(4), 513–541.

Parks, T. E. (1965). Post-retinal visual storage. *The American Journal of Psychology*, *78*, 145–147.

Plateau, J. A. F. (1836). Anorthoscop. *Bulletin de l'Acade´mie de Bruxelles*, *3*, 364.

Rock, I., Halper, F., DiVita, J., & Wheeler, D. (1987). Eye movement as a cue to figure motion in anorthoscopic perception. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 344–352.

Scholl, B. J., & Nevarez, H. G. (2002, November). Why so slow? The role of speed discontinuities in maintaining object persistence through occlusion. Paper presented at the annual meeting of the Psychonomic Society, Kansas City, MO.

Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology*, *38*, 259–290.

Shioiri, S., & Cavanagh, P. (1992). Visual persistence of figures defined by relative motion. *Vision Research*, *32*, 943–951.

Sohmiya, T., & Sohmiya, K. (1994). What is a crucial determinant in anorthoscopic perception? *Perceptual & Motor Skills*, *78*, 987–998.

Sperling, G. (1960). The information available in brief visual presentation. *Psychological Monographs*, *74*(11), 1–29.

Von Wright, J. (1968). Selection in visual immediate memory. *Quarterly Journal of Experimental Psychology*, *20*, 62–68.

Zöllner, F. (1862). Uber eine neue Art anorthoskopischer Zerrbilder (On a new type of anorthoscopic distorted image). *Annalen der Physic*, *117*, 477–484.

PART VII
Attention

Chapter 103
Motion-Induced Blindness

Yoram Bonneh

THE EFFECT

In a typical example (see Fig. VII.103-1a), a swirling cloud of blue dots is superimposed on three highly salient but stationary yellow spots. When one fixates anywhere on the display, after a while, one, two, or even all three spots simply disappear (Bonneh, Cooperman, & Sagi, 2001). The swirling cloud appears to remove the yellow spots from sight, even though the spots continue to stimulate the retina. The properties of this phenomenon can best be appreciated by viewing a dynamic demonstration (see http://www.scholarpedia.org/article/Motion_induced_blindness; www.weizmann.ac.il/~masagi/MIB/mib.html; www.michaelbach.de/ot/mot_mib/). Overall, the motion-induced blindness (MIB) phenomenon is quite robust. Although some observers fail to see the effect at the beginning, most tend to improve with practice, and some are even able to get small patterns to disappear almost indefinitely.

The MIB phenomenon can occur under more natural conditions than those described here and perhaps even in everyday life situations. Richard Brown, from the San Francisco science museum the Exploratorium, created a version of the phenomenon, which is illustrated in Figure VII.103-1b. Objects as salient as a silver watch or keys, which are placed on a round, transparent table, can disappear when a textured plate underneath the table is rotated. Interestingly, a strong disappearance effect occurs immediately after the onset of the rotation movement, thus demonstrating that a transient stimulus is especially effective in inducing disappearance. Shinsuke Shimojo, from the California Institute of Technology, used a mirror ball (as is used in discotheques) to create an optical flow field on the walls of a large, dark room. Watching a person standing a few meters away in this setting creates the illusion of a floating figure, followed by the complete disappearance of this figure. A third example of MIB is night driving, whereby the driver is exposed to the optical flow field of the road lights while fixating on the back lights of the car in front of him or her. These lights could possibly disappear especially when the driver is tired and does not move his or her eyes or blink much.

THE RELEVANCE OF THE ILLUSION

MIB has attracted much interest in relation to studying perceptual consciousness. To understand why, consider a situation in which a bright and salient spot suddenly disappears, as if "wiped out" by the moving pattern for several seconds. Perceptually, this is not different from erasing the spot on the screen, and. in fact, observers cannot tell the difference and are often convinced that the spot was erased. However, the bright spot reaches our eyes continuously, and its retinal response reaches our brain, creating a neural representation of an invisible object. This allows us to compare the neural responses under two physically identical conditions: one visible or conscious and the other invisible or subconscious. We can then investigate the neural representation that underlies visibility or, more generally, the neural correlate of conscious perception. For example, is consciousness represented in specific brain regions, perhaps excluding some low-level processing stages, or, alternatively, is it distributed in many regions, perhaps dynamically? Is it related to the firing rates of neurons or, alternatively, to specific temporal patterns of firing or synchrony? Other open questions are whether the representation of the invisible is accessible to interactions with visible stimuli and whether it retains all its properties compared with the representation of the visible. This approach of investigating the neural correlates of consciousness by studying visual perception has been advocated and pursued by Koch and Crick (2003).

VARIANTS AND PARAMETERS

MIB is typically observed and measured by continuously presenting a static *target* and a rotating *mask* (see Fig. VII.103-1). While viewing these stimuli, observers report events of disappearance and reappearance of a small target pattern by pressing buttons. This procedure yields three main parameters: (a) the proportion of time in which the disappearing target is invisible, (b) the rate of disappearance events (the number per minute), and (c) the average duration of the invisibility periods. These parameters are

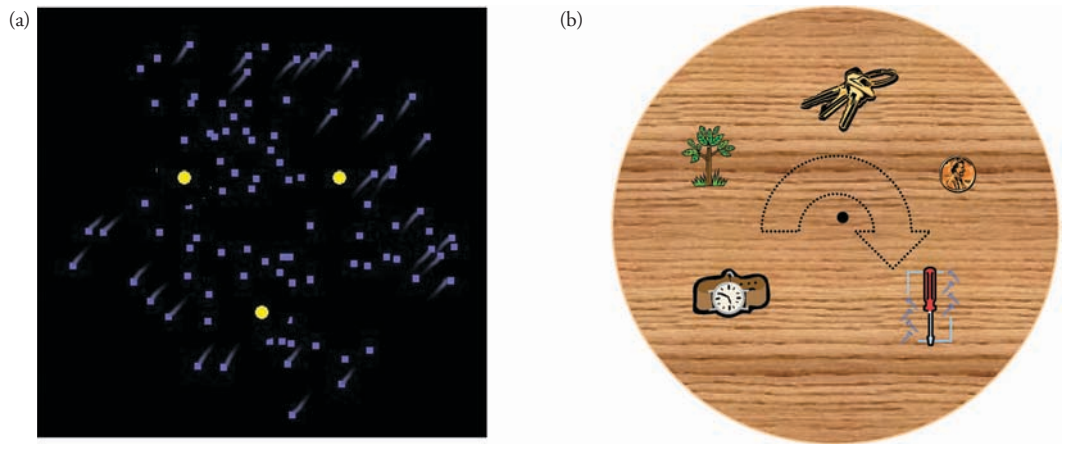


Figure VII.103-1. Typical stimuli that produce MBI. (a) A snapshot of a dynamic computer display in which the blue dots produce a three-dimensional sphere that rotates around, whereas the yellow dots are static. With steady fixation, observers report the disappearance of one or more yellow dots for periods of a few seconds at a time. (b) Illustration of a physical setup that produces MIB. Several salient objects are placed on a static transparent glass surface, under which a textured plate can be rotated. When the textured surface is set in motion while observers fixate at the center, the objects often disappear. Illustrations adapted from Bonneh, Cooperman, and Sagi (2001) in (a) and are based on a demonstration made by Richard Brown in San-Francisco Exploratorium in (b).

not independent, and two parameters can determine the third. For example, one can have a 50% disappearance with one 30 s invisibility period (rate = 1 per min), or 30 periods of 1 s (rate = 30 per min). Disappearance can reach 40% of the time and more, with a typical average invisibility period of 1 to 3 s. The duration and frequency of disappearance are determined by the properties of the target and the surrounding moving mask (Bonneh et al., 2001).

Target Size

Target size should generally be small, but patterns as large as a thumb at arm's-length distance can also disappear, depending on their distance from the fixation point.

Target Position

More disappearance occurs away from fixation and rarely or never at the fixation point. Targets with large eccentricity (e.g. above 4°), especially those with low contrast, could disappear without any mask (this is known as the Troxler effect). Interestingly, disappearance is not uniform across space. Patterns in the upper-left quadrant of the visual field tend to disappear more for most observers.

Target Movement and Flicker

Target movement and flicker usually reduce disappearance compared with static targets, but a dynamic target can also disappear. Most strikingly, a slowly moving peripheral target can disappear in one quadrant and reappear in another. Flickering targets disappear especially at fast rates (e.g., 3 Hz flickering dots disappear 25% of the time).

Spatial Layout of Targets

Parts forming a group (Gestalt) according to the principles of perceptual organization tend to disappear together, and parts that appear to belong to different objects tend to disappear in alternation (Bonneh et al., 2001).

Target Luminance and Contrast

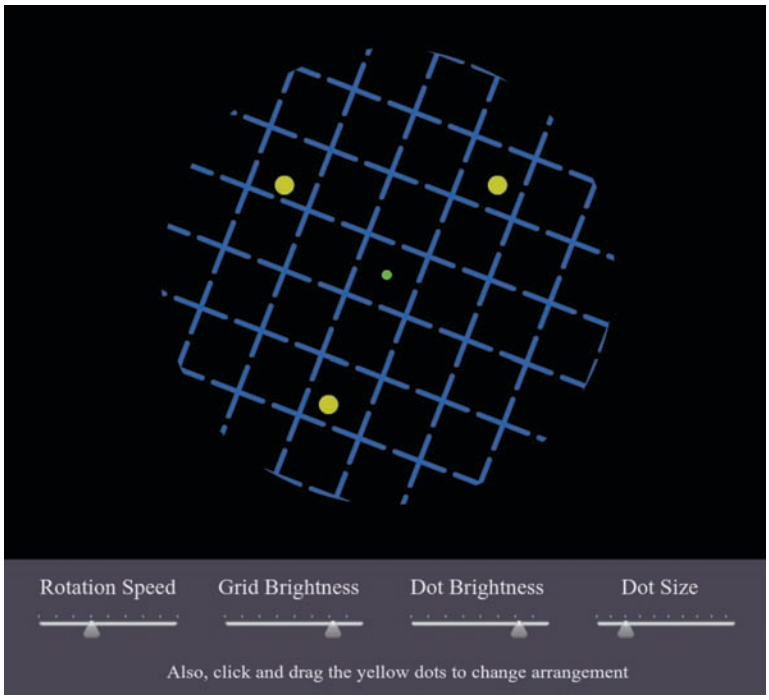
For a small, bright patch on a black background, the brighter the patch is, the more frequently it will disappear (Bonneh et al., 2001). This paradoxical finding is in disagreement with the Troxler effect (peripheral fading without a mask, see also later discussion).

Mask

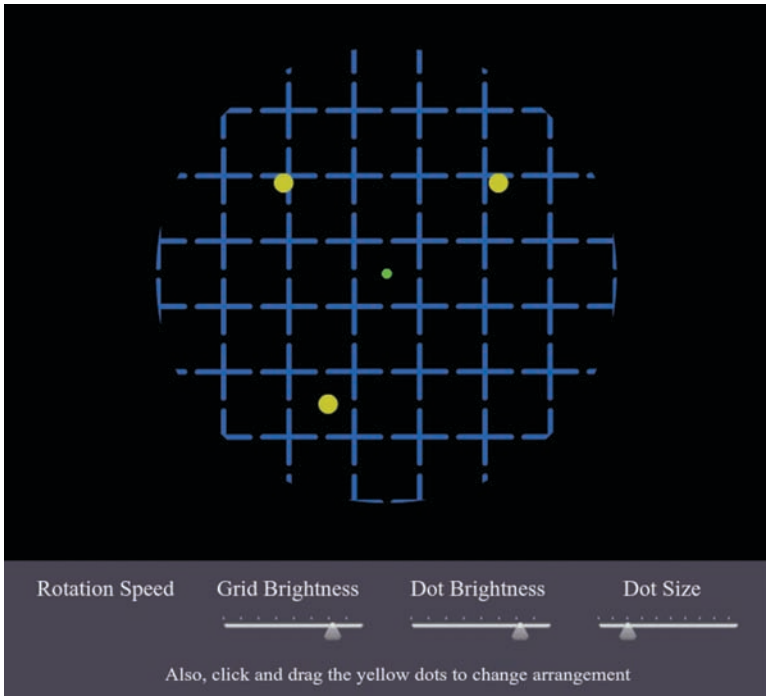
The type of moving mask used affects the disappearance, but many types of masks can induce some disappearance as long as the mask surrounds the target. This includes a rotating pattern on a two-dimensional surface (e.g., as in Figs. VII.103-1a and VII.103-1b), and a random dot three-dimensional sphere (also used in the original study, Bonneh et al., 2001), as well as random incoherent motion and local surrounding flicker (Kawabe & Miura, 2007). The number of dots in the mask need not be large. The mask contrast could also be very low as long as the mask is visible. The mask need not be close to the target, and the target could be surrounded by an empty "protection zone" (as in Fig. VII.103-1) of a few degrees and could readily disappear (Bonneh et al., 2001). More disappearance occurs when the mask is presented in front of the target (e.g., via binocular stereo disparity); this is interpreted as occlusion (Graf, Adams, & Lages, 2002).

Eye Movements

Abrupt saccades or microsaccades typically terminate the disappearance, although it can often survive small eye movements around the fixation point (Bonneh et al., 2010).



Interactive Figure VII.103-1. Interactive demonstration for Motion Induced Blindness. Fixate on the green dot in center of the display and one or more of the yellow dots will seem to disappear for seconds at a time. The sliders control rotation speed, grid brightness, dot brightness and dot size. The display also allows observers to also drag the yellow dots to new positions. (Demonstration adapted from Bonneh, Cooperman, and Sagi [2001].)



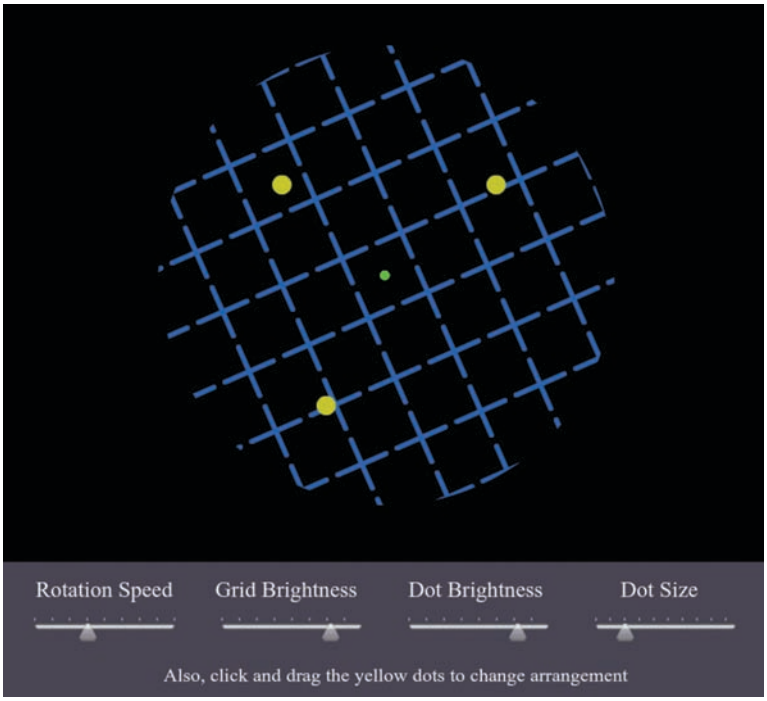
Interactive Figure VII.103-2. See caption for Interactive Figure VII.103-1.

DISCUSSION

Next we discuss different explanations and mechanisms possibly involved in MIB, as well as related phenomena.

Related Phenomena

There are several other related phenomena in which stimuli are not registered in conscious perception and thus become invisible (Kim & Blake, 2005). Such invisibility could be explicit (seeing it disappear) or implicit (failing to notice it). The conditions for implicit invisibility include a brief presentation, low contrast, a noisy background, and masking. Under these conditions, unlike MIB, stimuli do not disappear, but their sensory representation is degraded or the processing is disrupted. Implicit invisibility also occurs when stimuli escape our focus of attention, such as in change blindness, inattentional blindness, and the attentional blink (see Kim & Blake, 2005, for a



Interactive Figure VII.103-3.
See caption for Interactive Figure VII.103-1.

review). MIB belongs to a different class of phenomena in which the lack of perception has an explicit nature of "visual disappearance," as if stimuli are erased in front of the observers' eyes. This class of phenomena includes binocular rivalry, in which dissimilar patterns presented to different eyes disappear in alternation; Troxler fading, in which low-contrast peripheral stimuli disappear under strict fixation; and flash-suppression, in which disappearance in one eye is induced by a flash presented to the other. It also includes meta-contrast masking and perceptual filling-in of "artificial scotomas," in which static peripheral patches disappear when "filled-in" by dynamic surrounding noise (Ramachandran & Gregory, 1991). In contrast with the latter, MIB occurs with large empty "protection zones" that do not allow filling-in by the dynamic mask (Fig. VII.103-1). Another important feature of MIB is the stochastic alternation between the two perceptual states of visible and invisible. This links MIB to other multistable perception phenomena (Blake & Logothetis, 2002).

MIB and Retinal Stabilization

When an image is stabilized on the retina (e.g., when it is set to move with the eye), it fades and disappears similarly to that observed in MIB (Ditchburn & Ginsborg, 1952), a phenomenon well explained in terms of photoreceptor adaptation in the retina. If the moving mask alters or reduces the fixational eye movements, it may induce image stabilization, enhanced adaptation, and consequently fading. However, this is not the case, as revealed by a careful investigation of fixational eye movements in MIB (Bonneh et al., 2010). First, the pattern of microsaccades, which are the main component of fixational eye movements, does not change significantly in the presence of a rotating mask. Second, microsaccades, similar to any other transient visual stimulation, tend to trigger reappearance; however, they have primarily a modulatory effect and are neither necessary nor sufficient to account for MIB (Bonneh et al., 2010).

MIB and Attention

One early explanation of MIB suggests that it reflects spontaneous attention shifts induced by the moving mask (Bonneh et al., 2001), an idea consistent with the finding of competitive (anticorrelated) functional magnetic resonance imaging responses measured in dorsal and ventral extrastriate cortex during MIB (Donner, Sagi, Bonneh, & Heeger, 2008). Psychophysically, it has been found that withdrawing attention from both the target and the mask by a demanding task impedes the bistable process and prolongs the invisible periods (Scholvinck & Rees, 2009). On the other hand, directing spatial attention to the target increases the probability and duration of its disappearance (Scholvinck & Rees, 2009), similar to that observed with an increase in target contrast (Bonneh, Donner, Cooperman, Heeger, & Sagi, 2014; Bonneh et al., 2001).

NEURAL MECHANISMS UNDERLYING MIB

The explicit nature of target disappearance and its spatial specificity (multiple targets at different locations disappear in an uncorrelated fashion) suggest that MIB involves low-level, local sensory mechanisms, such as those in the primary visual cortex or even in the retina. This idea of local

low-level suppression is further supported by evidence for the involvement of adaptation (Gorea & Caetta, 2009), filling-in (Hsu, Yeh, & Kramer, 2006), and motion streak suppression (Wallis & Arnold, 2009) in MIB. However, explanations based entirely on early perceptual suppression do not seem sufficient for several reasons. First, a slowly moving pattern can disappear at one position and reappear after a few seconds on the other side of the screen, arguing against adaptation of localized neuronal populations being the sole mechanism involved. Second, as previously noted, the moving mask can be quite far (a few degrees of visual angle) from the target and still induce target disappearance, which rules out a local masking mechanism based on lateral connections in the retina. Third, the brighter the target, the more frequently it disappears, which rules out a simple "gain-control" mechanism (discarding low-contrast stimuli in a cluttered environment) as the basis of MIB (Bonneh et al., 2001).

Another line of evidence focuses on the fate of the invisible stimuli and the type of their residual subconscious processing. According to this viewpoint, if a visual object is completely erased at a certain level of a hierarchical process, then it will not affect higher-level processing and its disappearance properties will not be affected by these higher levels, at least not during the invisible periods. Importantly, it was shown that stimuli rendered invisible by MIB retain their capacity to produce orientation-selective adaptation (Montaser-Kouhsari, Moradi, Zandvakili, & Esteky, 2004) and can induce negative afterimages (Hofstoetter, Koch, & Kiper, 2004), suggesting a cortical representation of the invisible at these processing levels. Moreover, stimuli that are perceptually grouped into objects or Gestalts, such as collinear line segments or proximal dots, tend to disappear and reappear in synchrony and less frequently, rather than disappear and reappear independently (Bonneh et al., 2001). When the configuration is changed during the invisible stage, its reappearance depends on the new configuration, with good Gestalts reappearing in unity and bad Gestalts in parts (Mitroff & Scholl, 2005). This suggests that a representation of the invisible objects is available at the level of Gestalt grouping. Finally, the three-dimensional segmentation of the mask relative to the target has a significant effect on disappearance, with more disappearance observed when the mask is presented in front of the target and is interpreted as occluding it (Graf et al., 2002). Taken together, these findings of high-level processing of the invisible stimuli are inconsistent with explanations based entirely on early low-level suppression.

The evidence for both low-level and high-level mechanisms involved in MIB discussed earlier calls for explanations based on neural interactions between different levels. According to one hypothesis, the spontaneous target disappearance in MIB results from competitive interactions between distinct populations of cortical neurons processing the target and the mask (Donner et al., 2008; Keysers & Perrett, 2002). Whereas the static target is processed by the ventral pathway, the moving mask is processed primarily by the dorsal pathway. This hypothesis is supported by functional magnetic resonance imaging findings of opposite modulations in ventral and dorsal pathways during a target's illusory disappearance: the response to the target decreases in the ventral pathway (such as area V4), and the response to the mask increases in the dorsal pathway (such as area V5/MT; Donner et al., 2008). However, these opposite modulations do not occur during a physical "replay" of the target's disappearance, suggesting that they may play a causal role in MIB.

Why Do Salient Patterns Disappear in MIB?

It is often debated whether MIB reflects a failure of visual processing due to some limitation of the system or, alternatively, to a functional mechanism that discards suspected visual artifacts (New & Scholl, 2008). The evidence reviewed so far suggests that these two opposing views are not necessarily contradictory or mutually exclusive and can be reconciled assuming that different levels of processing are involved. At the sensory level, the limited transient stimulation of the targets may cause a decay of the response due to adaptation mechanisms. This occurs all the time, even without the surrounding motion, but the system normally knows not to interpret this decay as physical disappearance and can fill-in the missing representation over time. However, when there is a perceptual hint that suggests otherwise, then the system may interpret the sensory decay as real disappearance. Thus MIB results from a sensory-level process, an interpretation, or a decision-level process and their interactions.

REFERENCES

Blake, R., & Logothetis, N. K. (2002). Visual competition. *Nature Reviews Neuroscience*, *3*(1), 13–21.

Bonneh, Y. S., Cooperman, A., & Sagi, D. (2001). Motion-induced blindness in normal observers. *Nature*, *411*(6839), 798–801.

Bonneh, Y. S., Donner, T. H., Cooperman, A., Heeger, D. J., & Sagi, D. (2014). Motion-induced blindness and Troxler fading: Similarities and differences. *PLoS ONE*, *9*(3), e92894.

Bonneh, Y. S., Donner, T. H., Sagi, D., Fried, M., Cooperman, A., Heeger, D. J., & Arieli, A. (2010). Motion-induced blindness and microsaccades: Cause and effect. *Journal of Vision*, *10*(14), 22.

Crick, F., & Koch, C. (2003). A framework for consciousness. *Nature Neuroscience*, *6*(2), 119–126.

Ditchburn, R. W., & Ginsborg, B. L. (1952). Vision with a stabilized retinal image. *Nature*, *170*, 36–37.

Donner, T. H., Sagi, D., Bonneh, Y. S., & Heeger, D. J. (2008). Opposite neural signatures of motion-induced blindness in human dorsal and ventral visual cortex. *Journal of Neuroscience*, *28*(41), 10298–10310.

Gorea, A., & Caetta, F. (2009). Adaptation and prolonged inhibition as a main cause of motion-induced blindness. *Journal of Vision, 9*(6):16, 1–17.

Graf, E. W., Adams, W. J., & Lages, M. (2002). Modulating motion-induced blindness with depth ordering and surface completion. *Vision Research*, *42*(25), 2731–2735.

Hofstoetter, C., Koch, C., & Kiper, D. C. (2004). Motion-induced blindness does not affect the formation of negative afterimages. *Consciousness and Cognition*, *13*(4), 691–708.

Hsu, L. C., Yeh, S. L., & Kramer, P. (2006). A common mechanism for perceptual filling-in and motion-induced blindness. *Vision Research*, *46*(12), 1973–1981.

Kawabe, T., & Miura, K. (2007). Subjective disappearance of a target by flickering flankers. *Vision Research*, *47*(7), 913–918.

Keysers, C., & Perrett, D. I. (2002). Visual masking and RSVP reveal neural competition. *Trends in Cognitive Sciences*, *6*(3), 120–125.

Kim, C. Y., & Blake, R. (2005). Psychophysical magic: Rendering the visible "invisible." *Trends in Cognitive Sciences, 9*(8), 381–388.

Mitroff, S. R., & Scholl, B. J. (2005). Forming and updating object representations without awareness: Evidence from motion-induced blindness. *Vision Research. 45*(8), 961–967.

Montaser-Kouhsari, L., Moradi, F., Zandvakili, A., & Esteky, H. (2004). Orientation-selective adaptation during motion-induced blindness. *Perception*, *33*(2), 249–254.

New, J. J., & Scholl, B. J. (2008). "Perceptual scotomas": A functional account of motion-induced blindness. *Psychological Science*, *19*(7), 653–659.

Ramachandran, V. S., & Gregory, R. L. (1991). Perceptual filling in of artificially induced scotomas in human vision. *Nature*, *350*(6320), 699–702.

Scholvinck, M. L., & Rees, G. (2009). Attentional influences on the dynamics of motion-induced blindness. *Journal of Vision*, *9*(1):38, 1–9.

Wallis, T. S., & Arnold, D. H. (2009). Motion-induced blindness and motion streak suppression. *Current Biology*, *19*(4), 325–329.

Chapter 104
Inattentional Blindness and the *Illusion* of Attention

Daniel J. Simons

THE EFFECT

Take a look at Figure VII.104-1. The figure shows 3 frames from what is perhaps the best-known demonstration of a phenomenon known as "inattentional blindness": when people focus their attention on some aspect of a display, they often fail to notice an unexpected object or event. In this case, when tasked with counting the number of times the players wearing white shirts pass a basketball, about 50% of observers fail to notice a person in a gorilla suit who walks into the scene, stops midway through to thump her chest at the camera, and exits the other side.

WHAT MAKES INATTENTIONAL BLINDNESS INTERESTING

The finding that people miss the gorilla follows naturally from earlier work and theorizing (Neisser, 1979), but people still find it counterintuitive: Ninety percent of people predicted that they would notice the gorilla (Levin & Angelone, 2008), and in a national survey of American adults, 78% agreed that they generally notice unexpected events even when they are paying attention to something else (Simons & Chabris, 2011). This counterintuitiveness makes the invisible gorilla effect compelling. People are surprised when they miss the gorilla, occasionally accusing the experimenters of switching the tape after spotting it on a second viewing (Simons & Chabris, 1999).

Unlike most of the visual illusions discussed in this book, inattentional blindness is not a visual illusion but a cognitive one. We see the gorilla accurately when we are not engaged in counting passes, so the display alone does not induce inattentional blindness. Most illusions illustrate the shortcuts used by the visual system to make sense of the visual world, highlighting ways that our perceptual interpretation of a stimulus differs from the ground truth of the stimulus itself. Visual illusions reveal a nonisomorphic relationship between the world and what we see; we see things in a way that is different from the reality of the physical stimulus. An anamorphic painting or a space like the Ames room represent simple illusions. In those demonstrations, what we see is entirely consistent with the retinal image we receive from the stimulus, but it is inconsistent with the physical reality of the scene. With more sophisticated illusions, our visual system interprets the retinal stimulation as something that does not exist in the world. When a static image appears to move or when straight lines appear curved, we know that the illusion taps visual mechanisms that work well under most real-world viewing conditions but that break down for that particular stimulus. Sophisticated illusions of that sort "break" the visual system by presenting an atypical stimulus and showing how the visual system makes sense of it. In so doing, illusionists reveal the mechanisms that operate effectively when viewing more typical, real-world scenes. In effect, most optical and visual illusions present an "accidental" view of the world, one that is not representative of normal viewing conditions but that reveals the assumptions made by our visual system about such normal viewing conditions.

Inattentional blindness, in contrast, is not driven by a mistaken extrapolation from the two-dimensional retinal images to a three-dimensional world or from a misinterpretation of motion over time. It is not a failure to see something as it is, but a failure to see it altogether. In that sense, inattentional blindness is a more fundamental illusion in that it reveals a more complete discrepancy between the retinal sensation and conscious perception (see Changizi, 2010, for a similar assessment). The illusion lies not in the assumptions built into our visual system but in the beliefs we hold about the workings of our own minds. We see what we expect to see rather than what is there.

The phenomenon of inattentional blindness is interesting precisely because it is surprising and counterintuitive. No one would be surprised if observers failed to notice a camouflaged gorilla, but they believe that distinctive, unexpected objects will automatically capture their attention. The *illusion* is our mistaken belief that we will consciously see whatever confronts our visual system.

VARIANTS AND RELEVANT PARAMETERS

Inattentional blindness is remarkably easy to induce in viewers. At least some observers are likely to miss an unexpected object provided that (a) they are performing an attention-demanding task and (b) the unexpected object



Figure VII.104-1. Three frames from a video used by Simons and Chabris (1999) to study in attentional blindness. The video depicts the appearance of a woman in a gorilla suit walking through a ball-passing game.

or event is unrelated to that task or the attended items. Several factors appear to influence the rate of noticing. First, the more demanding the primary task, the less likely people are to notice the gorilla. Across a range of conditions, when observers keep separate tallies of aerial passes and bounce passes, noticing rates dropped by an average of approximately 20% (Simons & Chabris, 1999). Second, the similarity of the unexpected object to the attended and ignored items influences detection. For other variants of the gorilla video in which the gorilla does not stop to thump its chest, noticing rates are approximately 30% higher when counting passes by the players wearing black than when counting passes by the players wearing white (Simons & Chabris, 1999). Presumably people are more likely to notice the gorilla in that condition because its color more closely matches the attended black-shirted players. The role of similarity between the unexpected item and the attended items (and dissimilarity from the ignored items) was confirmed using better-controlled, computer-based inattentional blindness tasks (see Most et al., 2000).

Given that the phenomenon of inattentional blindness is driven, at least in part, by expectations and focused attention, the role of similarity and the attention demands of the primary task are unsurprising. More surprising is the lack of an effect of fixation. Observers viewing the gorilla video fixated the gorilla for an average of 1 s, with no difference in fixation times between those who saw the gorilla and those who missed it (Memmert, 2006; see also Beanland & Pammer, 2010).

Other factors might influence noticing rates as well. Experienced club basketball players were somewhat more likely to notice the gorilla than were other team-sport athletes (Memmert, 2006), perhaps because they found it easier to monitor the passes. However, in other inattentional blindness tasks, individual differences in monitoring performance did not predict noticing rates (Simons & Jensen, 2009), so the ability to perform the primary task does not appear to predict noticing rates in general.

One factor that does predict noticing of the gorilla is familiarity with the gorilla video. When observers know to look for a gorilla, they consistently see it. But observers who know to look for a gorilla and successfully spot it still miss other unexpected events in the same video (Simons, 2010).

DISCUSSION OF RESEARCH AND THEORIES

The term "inattentional blindness," brought to the attention of the broader vision community in Mack and Rock's (1998) book by that name, derives from the idea that people fail to see the unexpected event because their attention is focused elsewhere. That is, the blindness is caused by inattention. Mack and Rock used simpler displays in which observers judged which of two arms of a briefly flashed cross was longer. On a critical trial, in addition to the cross, another shape (e.g., a colored square, a word, an icon, etc.) appeared unexpectedly (see Figure VII.104-2). Depending on the nature of the unexpected object and the particular variant of the task, noticing ranged from approximately 25% to 75%.

Mack and Rock (1998) gave the effect its modern name and revived interest in the role of expectations in the detection of unexpected events, but the phenomenon itself has a much longer history. Analogous effects for auditory perception had been described in the literature on dichotic listening in the 1950s and 1960s (Cherry, 1953; Moray, 1959; Treisman, 1964). For centuries, magicians have tapped the phenomenon of inattentional blindness, using the limits of attention to hide their methods in plain sight (e.g., Fitzkee, 1945/2009). The effect was demonstrated informally in psychology by Hugo Münsterberg (1908) more than 100 years ago. He asked a classroom of students to focus on a color wheel, and many failed to notice a series of actions completed with his other hand. The first systematic studies of inattentional blindness also come from an unexpected literature: the study of ghosts. Parapsychologist Anthony Cornell (1959a, 1959b) conducted a series of studies in the 1950s in which he dressed as a ghost and made appearances on a college campus or on the stage in front of a movie. His goal was to recreate a paranormal experience in a way that could be replicated more systematically in a lab (i.e., without having to wait for the "true" appearance of a ghost in order to gauge how people react). He was disappointed—few people noticed him (although he reported that several cows followed him around campus; see Simons, 2011, for a more detailed description of these entertaining studies.)

In the 1970s, Ulric Neisser and his colleagues created a visual analog of a dichotic listening task, a method he termed "selective looking" (Neisser, 1979; Neisser &

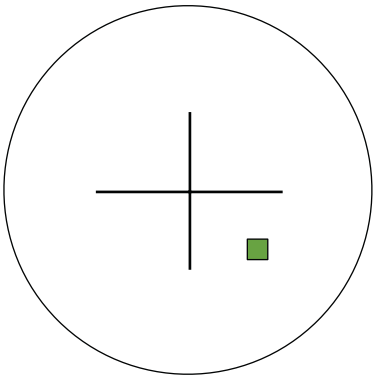


Figure VII.104-2. Recreation of a stimulus from the inattentional blindness task used by Mack and Rock (1998).

Becklen, 1975). In most of his tasks, observers viewed visually overlapping events and monitored just one of the events. For example, they might monitor a hand-slapping game that was visually superimposed (with partial transparency) on a video of people passing a ball. As observers counted passes (or hand slaps), something unexpected happened in the ignored event, and people often missed it. In one variant of the studies—the one that inspired the invisible gorilla study—people monitored players passing basketballs and often failed to notice a woman with an open umbrella who walked through the scene (Neisser, 1979; see Figure VII.104-3).

The key to Neisser's studies was that all of the players as well as the unexpected object physically occupied the same space in the scene; they were filmed separately and then superimposed. This partial transparency allowed Neisser to test whether attention selects regions of space or objects. If attention operates like a spotlight, highlighting a region of space and bringing to awareness anything within the beam, then observers should notice unexpected events that occupy the same space as other attended objects. However, if attention selects objects, then the umbrella woman could walk through the same space as the attended objects without being noticed. His finding of inattentional blindness was among the first powerful demonstrations of such object-based selective attention.

Although Neisser's work was well known at the time, it was not widely integrated into theories of attention, perhaps because it was ahead its time (see Most, Scholl, Clifford, & Simons, 2005, for discussion). Anecdotally at least, some of my colleagues dismissed the failure to notice the umbrella woman precisely because of the partial transparency of the displays. The black-and-white, grainy, partially transparent footage gave the displays an odd, ghostly appearance and made the unexpected event somewhat subtle. And that subtlety allowed researchers to discount the idea that such failures of awareness might occur for more vivid, distinct, and visible events. In other words, people could generate a plausible, intuitive reason why they might miss something unexpected. The invisible gorilla study, inspired directly by Neisser and his studies, was an attempt to replicate and extend his original work and to test that intuition by using a fully visible event. The study bore out the theoretical position Neisser espoused, showing that unexpected events, even distinctive ones, do not automatically grab attention.



Figure VII.104-3. Still image from the video described by Neisser (1979) to study selective looking. The video shows a woman with an open umbrella walking through a ball-passing game.

More recent studies have built on Neisser's work and the invisible gorilla study to explore the role of expectations (e.g., Most et al., 2005), the effects of similarity (e.g., Most et al., 2000), individual differences in noticing (e.g., Bredemeier & Simons, 2012; Richards, Hannon, & Derakshan, 2010; Seegmiller, Watson, & Strayer, 2011; Simons & Jensen, 2009), and the real-world ramifications of such failures of awareness (Chabris, Weinberger, Fontaine, & Simons, 2011; Hyman, Boss, Wise, McKenzie, & Caggiano, 2010). (For a relatively recent overview of inattentional blindness and the related phenomenon of change blindness, see Jensen, Yao, Street, & Simons, 2011.)

REFERENCES

Beanland, V., & Pammer, K. (2010). Looking without seeing or seeing without looking? Eye movements in sustained inattentional blindness. *Vision Research*, *50*, 977–988.

Bredemeier, K., & Simons, D. J. (2012). Working memory and inattentional blindness. *Psychonomic Bulletin & Review*, *19*, 239–244.

Chabris, C. F., Weinberger, A., Fontaine, M., & Simons, D. J. (2011). You do not talk about fight club if you do not notice fight club: Inattentional blindness for a simulated real-world assault. *i-Perception*, *2*, 150–153.

Changizi, M. (2010). Levels of real world wizardry: A review of the invisible gorilla by Chris Chabris and Dan Simons. http://goo.gl/tpzst

Cherry, E. C. (1953). Some experiments upon the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America, 25*, 975–979.

Cornell, A. D. (1959a). An experiment in apparitional observation and findings. *Journal of the Society for Psychical Research*, *40*(701), 120–124.

Cornell, A. D. (1959b). Further experiments in apparitional observation. *Journal of the Society for Psychical Research*, *40*(706), 409–418.

Fitzkee, D. (2009). *Magic by misdirection: An explanation of the psychology of deception*. Provo, UT: Magic Box Productions. (Original work published 1945)

Hyman, I. E. Jr., Boss, S. M., Wise, B. M., McKenzie, K. E., & Caggiano, J. M. (2010). Did you see the unicycling clown? Inattentional blindness while walking and talking on a cell phone. *Applied Cognitive Psychology*, *24*, 597–607.

Jensen, M. W., Yao, R., Street, W. N., & Simons, D. J. (2011). Change blindness and inattentional blindness. *WIREs Cognitive Science*, *2*, 529–546.

Levin D. T., & Angelone, B. L. (2008). The visual metacognition questionnaire: A measure of intuitions about vision. *The American Journal of Psychology*, *121*, 451–472.

Mack, A., & Rock, I. (1998). *Inattentional blindness*. Cambridge, MA: MIT Press.

Memmert, D. (2006). The effects of eye movements, age, and expertise on inattentional blindness. *Consciousness and Cognition*, *15*(3), *620–627*.

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology, 11*, 56–60.

Most, S. B., Scholl, B. J., Clifford, E. R., & Simons, D. J. (2005). What you see is what you set: Sustained inattentional blindness and the capture of awareness. *Psychological Review*, *112*(1), 217–242.

Most, S. B., Simons, D. J., Scholl, B. J., Jimenez, R., Clifford, E., & Chabris, C. F. (2000). How not to be seen: The contribution of similarity and selective ignoring to sustained inattentional blindness. *Psychological Science*, *12*, 9–17.

Münsterberg, H. (1908). *On the witness stand: Essays on psychology and crime*. New York: Clark Boardman.

Neisser, U. (1979). The control of information pickup in selective looking. In A. D. Pick (Ed.), *Perception and its development: A tribute to Eleanor J. Gibson* (pp. 201–219). Hillsdale, NJ: Lawrence Erlbaum.

Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive Psychology*, *7*, 480–494.

Richards, A., Hannon, E., & Derakshan, N. (2010). Predicting and manipulating the incidence of inattentional blindness. *Psychological Research*, *74*, 513–523.

Seegmiller, J. K., Watson, J. M., & Strayer, D. L. (2011). Individual differences in susceptibility to inattentional blindness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 785–791.

Simons, D. J. (2010). Monkeying around with the gorillas in our midst: Familiarity with an inattentional-blindness task does not improve the detection of unexpected events. *i-Perception*, *1*, 3–6.

Simons, D. J. (2011). Ghost busters, parapsychology, and the first study of inattentional blindness. http://goo.gl/ivuZQ

Simons, D. J., & Chabris, C. F. (1999) Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*, *28*, 1059–1074.

Simons, D. J., & Chabris, C. F. (2011). What people believe about how memory works: A representative survey of the U.S. population. *PLoS ONE*, *6*(8), e22757.

Simons, D. J., & Jensen, M. S. (2009). The effects of individual differences and task difficulty on inattentional blindness. *Psychonomic Bulletin & Review*, *16*(2), 398–403.

Treisman, A. (1964). Monitoring and storage of irrelevant messages in selective attention. *Journal of Verbal Learning and Verbal Behavior*, *3*, 449–459.

PART VIII
Binocular Vision/Stereopsis

Chapter 105
Binocular Rivalry
The Illusion of Disappearance

Randolph Blake

Visual illusions entail dissociations between some property of an object or an event (e.g., the amount of light reflected from an object's surface) and perceptual appearance of that object or event (e.g., surface lightness). Visual appearance, in other words, departs from objective reality in some intriguing, often nonintuitive way. Given that characterization of illusions, imagine this scenario: something one is looking at suddenly and unexpectedly vanishes from visual awareness, and yet nothing has happened in the external environment that might cause it to disappear. This beguiling, magical event would surely qualify as a compelling visual illusion indeed. In fact, this is precisely what happens during binocular rivalry,[1] and that is why this fascinating phenomenon readily merits inclusion in this compendium.

Binocular rivalry is instigated when dissimilar visual stimuli are imaged on corresponding areas of the two eyes. As illustrated in Figure VIII.105-1, those dissimilarities can arise from differences in form (both simple and complex), color, or direction of motion. During normal viewing, of course, our two eyes are looking at the same objects within a common visual scene, insuring that the foveae of the two eyes receive matched monocular images. But in the laboratory we can fool the binocular visual system by using an optical device (e.g., a stereoscope) to present entirely different images or different movie animations to the left and right eyes. Under these conditions of dissimilar monocular stimulation, stable binocular single vision gives way to perceptual competition between the two pictures or animations: rather than perceiving an amalgam of the two dissimilar stimuli, we will experience alternations in perceptual dominance wherein one stimulus alone is perceived at any given moment, with the other stimulus suppressed from vision. These periods of monocular dominance can last for several seconds or more, and the transitions in dominance from one stimulus to the other are not abrupt but appear like waves of change as one stimulus sweeps the other one from awareness. If the two stimuli are relatively large, we may also experience mixed dominance in which bits and pieces of the two images appear spatially intermingled, forming a patchwork-like visual impression that itself changes with time. Fluctuations in perception tend to occur unpredictably, and the durations of dominance vary considerably during an extended viewing period (Fig. VIII.105-2), implying that switches in perceptual dominance during rivalry are not triggered by a periodic, clocklike signal. The most common view is that switches in dominance result from a combination of neural adaptation and internal neural noise (Brascamp, van Ee, Noest, Jacobs, & van Den Berg, 2006).

A few years ago, a particularly potent technique for inducing binocular rivalry was introduced to the field of vision by Tsuchiya and Koch (2005). Illustrated in Figure VIII.105-3, the technique entails presenting a sequence of richly contoured, Mondrian-like patterns to one eye at the rate of about 10 patterns/s while, at the same time, presenting to the other eye an ordinarily visible picture of anything ranging from a one-dimensional grating pattern to a familiar face or a natural scene. The dynamic Mondrian patterns can render the picture invisible for more than a minute at a time, and for this reason the technique has been dubbed "continuous flash suppression."

Like many other illusions, binocular rivalry cannot be overridden by actual knowledge of what we are looking at—sheer mental will power is powerless to completely halt alternations in rivalry (Blake, 1988). People can influence the amount of time that a given stimulus remains perceptually dominant, however, by carefully attending to that dominant stimulus to detect, for example, small, unpredictable changes in its shape (Chong, Tadin, & Blake, 2005). In a related vein, individuals trained in specific forms of meditation can slow, but not halt, their rivalry alternations (Carter, Presti, Callistemon, et al., 2005). Ingestion of certain drugs, too, can temporarily alter the rate at which perception switches during rivalry, with some drugs speeding up the switch rate (psychedelics such as ayahuasca; Frecska, White & Luna, 2003) and others slowing them down (alcohol; Donnelly & Miller, 1995). Within the general population, the average rate of alternations during rivalry varies widely among individuals (Carter & Pettigrew, 2003), and this trait appears to have a genetic component (Miller, Hansell, Ngo, et al., 2010). Binocular rivalry alternations also are influenced by sensory information conveyed by other modalities. For instance, sounds denoting a particular object (e.g., a bird singing) can bias dominance of a picture of that object (a bird) when it is engaged in

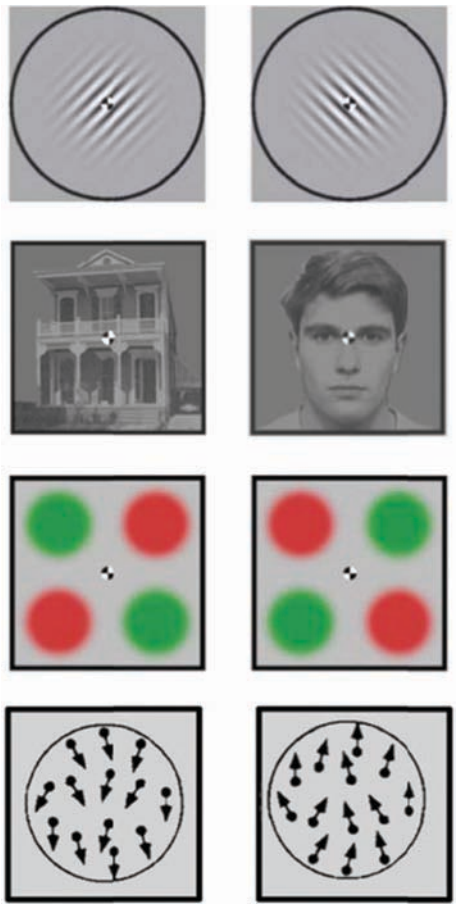


Figure VIII.105-1. Examples of the kinds of dichoptic displays (i.e., different stimuli viewed by the two eyes) that can trigger binocular rivalry. (a) Dichoptic orthogonal gratings. (b) Conflicting objects. (c) Mismatched colors. (d) Arrays of dots generally moving in opposite directions. Using any of these pairs except the bottom one, one may be able to experience binocular rivalry by crossing one's eyes until the two images appear aligned (with the left eye focused on the right image and the right eye focused on the left image). Alternatively, one can use a pair of red–green anaglyph glasses to view rivalry images at http://www.psy.vanderbilt.edu/faculty/blake/rivalry/BR.html. (House/face pair courtesy of Frank Tong.)

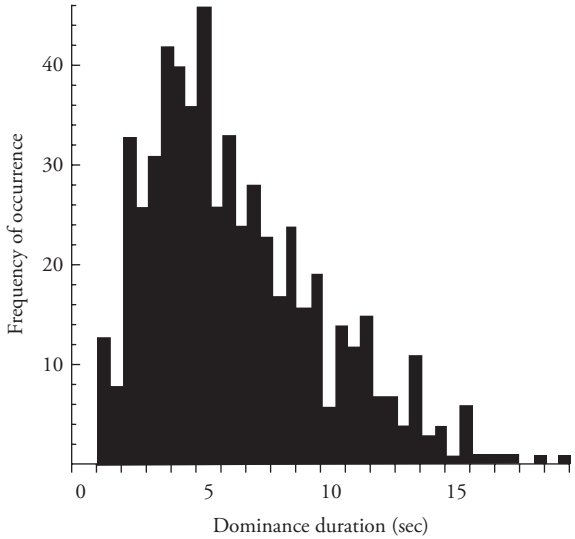


Figure VIII.105-2. Typical frequency histogram of dominance durations measured during extended periods of binocular rivalry. Durations vary irregularly from episode of exclusive dominance to the next, with only a few lasting for very brief durations and a few lasting for very long durations; the result is a distribution of durations exhibiting a notable rightward skew. (Data courtesy of Martin Lankheet, from Lankheet, 2006, and redrawn here with his permission and that of the Association for Research in Vision and Ophthalmology, which who holds the copyright.)

rivalry with another picture of another object unrelated to the sound (Chen, Yeh, & Spence, 2011). Similarly, smelling an odorant signifying, say, a rose will bias perception of an image of a rose when it is engaged in rivalry with a non-rose-like image (Zhou, Jiang, He, & Chen, 2010). Even touching a corrugated surface biases visual rivalry in favor of a picture of an object whose striations run in the direction of the surface being touched (Lunghi, Binda, & Morrone, 2010). The motor system also seems to be involved in the process, as evidenced by the fact that self-generated movement of the hand by a person viewing rivalry can influence the durations of dominance and suppression (Maruya, Yang, & Blake, 2007). These findings implicate the involvement in top-down influences on the processes governing dominance during binocular rivalry.

Returning to the characteristic of rivalry that qualifies it as an illusion, an ordinarily visible stimulus is erased from awareness for seconds at a time, even though that stimulus remains imaged on the retina—perceptually speaking, a suppressed stimulus temporarily ceases to exist. This aspect of rivalry makes it an effective tool for studying subliminal perception (Lin & He, 2009), allowing us to ask: "Does a stimulus erased from awareness retain some of its effectiveness to impact other aspects of vision?" The answer appears to be "no" when it comes to rivalry's effect on other visual illusions involving an alteration in an object's appearance consequent to the presence of neighboring visual objects—suppression of those neighbors abolishes the illusion (Meng, Remus, & Tong, 2005; Rao, 1977; Sobel & Blake, 2003; but see Wade, 1980). What about suppression's effect on the build-up of visual adaptation as gauged by the magnitude of the aftereffects produced by adaptation—does suppression forestall adaptation? Here the answer is "sometimes." Simple visual patterns such as oriented contours can still induce adaptation aftereffects, albeit more weakly, even when they are suppressed (Blake, Tadin, Sobel, et al., 2006), and the same is true for simple forms of visual motion (Maruya, Watanabe, & Watanabe, 2008). But aftereffects produced by adaptation to more complex patterns such as human faces are drastically attenuated (Yang, Hong, & Blake, 2010) or abolished completely by suppression (Moradi, Koch, & Shimojo, 2005), as is adaptation to complex forms of motion (Maruya, Watanabe, & Watanabe, 2008; Wiesenfelder & Blake, 1990). Some evidence implies that attention can be attracted to an emotionally charged picture (e.g., a nude human) even when that stimulus is suppressed from visual awareness (Jiang, Costello, Fangm, Huang, & He, 2006); words, on the other hand, do not appear to retain any effectiveness when fully suppressed from awareness (Kang, Blake, & Woodman, 2011). To be sure, these kinds of questions

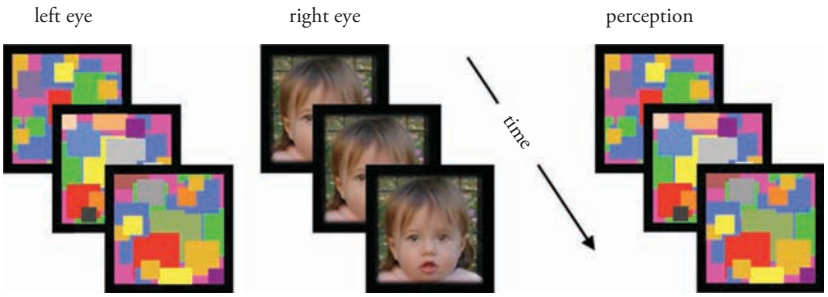


Figure VIII.105-3. Schematic of the continuous flash suppression technique. A rapidly changing sequence of Mondrian-like arrays presented to one eye will render another stimulus presented to the other eye completely invisible for many seconds at a time (even a stimulus as adorable as a young grandchild). The Mondrian patterns need not be colored; the crucial ingredients for producing continuous flash suppression are the rapid (~ 10 Hz), random changes in shapes, colors and lightness values of the components of the Mondrian. (Author's own creation.)

about subliminal perception can be tackled using other illusions such as motion-induced blindness (e.g., Montaser-Kouhsari, Moradi, Zandvakili, & Esteky, 2004), but none can rival binocular rivalry when it comes to erasing a wide variety of visual stimuli for relatively long periods of time. For that reason, those interested in visual processing outside of awareness have increasingly turned to binocular rivalry as a favored means for rendering an ordinarily visible stimulus utterly invisible (Kim & Blake, 2005).

Thus during binocular rivalry we are fooled into experiencing changes in visual perception over time; however, those changes are not happening in the world but instead are arising entirely from neural events within our brains. For this reason, binocular rivalry affords a powerful tool for studying neural events accompanying fluctuations in visual perception: at some stages of processing, neural responses should wax and wane in synchrony with fluctuations in dominance and suppression. The search for those neural concomitants of rivalry alternations has been underway for years now, using single-cell recording techniques in animals and brain-imaging techniques in humans (Tong, 2005). What has been discovered? We see rivalry-related modulations in neural activity throughout the visual hierarchy (Blake & Logothetis, 2002; Tong, Meng, & Blake., 2006), ranging from the thalamus, up through early visual cortex and into high-tier ventral areas, In addition, there are telltale signs of neural activity associated with rivalry transitions appearing within frontal and parietal areas (Knapen, Brascamp, Pearson, van Ee, & Blake, 2011; Lumer, Friston, & Rees, 1998). So it looks like the neural events underlying rivalry are widely distributed, with the impact of neural suppression seeming to grow deeper as one moves up the visual hierarchy (Nguyen, Freeman, & Alais, 2003). According to this idea, then, it is an oversimplification to speak of *the* locus of binocular rivalry, and for that reason older theories of rivalry (e.g., Blake, 1989) have been supplanted by theories (e.g., Hohwy, Roepstorff, & Friston, 2008) in which a confluence of bottom-up and top-down influences orchestrate the neural battle between competing stimuli viewed by the two eyes.

Aside from its fascinating, illusory character, binocular rivalry may bear importantly on other questions concerning vision. We know, for example, that rivalry dominance can become synchronized over large regions of the visual field, yielding global dominance of spatially separated object features (Kovács, Papathomas, Yang, & Feher, 1996). This implies that rivalry can tell us something about global perceptual organization. Binocular rivalry also offers an opportunity to learn about constructive processes shaping normal visual perception through attention, expectations, and other "top-down" processes. In this vein, rivalry has been singled out as a paradigm phenomenon for characterizing the inferential nature of vision within the context of predictive coding and Bayesian inference (Dayan, 1998). Moreover, binocular rivalry may be closely allied with binocular stereopsis, the perception of depth arising from perspective differences arising from the slightly different views of the world provided by our left and right eyes (Grossberg, Yazdanbakhsh, Cao, & Swaminathan, 2008). On a practical note, rivalry may have implications for the design and use of head-worn three-dimensional displays such as those used in commercial and military applications, for those devices can unwittingly induce binocular rivalry and, therefore, impair visual performance (Patterson, Winterbottom, Pierce, & Fox, 2007). In addition, rivalry may have relevance for understanding certain health-related problems. One such problem is the suppression of vision accompanying chronic eye misalignment, a condition that causes corresponding areas of the two eyes to receive discrepant inputs. Called strabismus, this condition can lead to a permanent, unilateral loss of form vision (von Noorden, 1990). Abnormalities in rivalry dynamics have also been implicated in some categories of psychopathology, including bipolar disorder and schizophrenia (Tononi & Edelman, 2000). Indeed, abnormal rivalry dynamics may denote a genetic propensity for schizophrenia (Frecska, White, Leonard et al., 2003).

Finally, it is worth noting that the illusory fluctuations in perception characterizing binocular rivalry are not limited to vision. Dissimilar auditory messages broadcast to the two ears compete for dominance, creating binaural rivalry (Brancucci & Tommasi, 2011). Dissimilar odorants introduced separately to the two nostrils induce binaural rivalry, wherein the odorants alternate in perceptual dominance (Zhou & Chen, 2008). Tactile rivalry occurs also, when sequences of brief vibrotactile stimulation are applied to a finger in a configuration that generates two competing perceptual interpretations of tactile motion (Carter, Konkle, Wang, Hayward, & Moore, 2008). In general, illusory fluctuations in perception seem to be the rule when sensory inputs imply conflicting interpretations, and the evidence favoring one interpretation over another is weak (Arnold, 2011).

ACKNOWLEDGMENTS

This chapter was prepared while the author was supported by the World Class University program through the Korea Science and Engineering Foundation funded by the Ministry of Education, Science and Technology (R31-10089).

NOTE

1. Binocular rivalry has a storied history dating back to the 16th century (Wade, 1998), but it was British physicist Sir Charles Wheatstone (1852) who first systematically studied binocular rivalry. Using his newly invented stereoscope for achieving independent stimulation of the two eyes, Wheatstone was able to examine in detail what happens when dissimilar visual stimuli are imaged on corresponding areas of the two eyes.

REFERENCES

Arnold, D. H. (2011). Why is binocular rivalry uncommon? Discrepant monocular images in the real world. *Frontiers in Human Neuroscience, 5*, 116. doi:10.3389/fnhum.2011.00116

Blake, R. (1988). Dichoptic reading: The role of meaning on binocular rivalry. *Perception & Psychophysics*, *44*, 133–141.

Blake, R. (1989). A neural theory of binocular rivalry. *Psychological Review*, *96*, 145–167.

Blake, R., & Logothetis, N. (2002). Visual competition. *Nature Reviews Neuroscience*, *3*, 13–23.

Blake, R., Tadin, D., Sobel, K., Chong, S. C., & Raissian, R. (2006). Strength of early visual adaptation depends on visual awareness. *Proceedings of the National Academy of Sciences USA*, *103*, 4783–4788.

Brancucci, A., & L. Tommasi (2011). Binaural rivalry: Dichotic listening as a tool for the investigation of the neural correlate of consciousness. *Brain and Cognition*, *76*, 218–224.

Brascamp, J. W., van Ee, R., Noest, A. J., Jacobs, R. H., & van Den Berg, A. V. (2006). The time course of binocular rivalry reveals a fundamental role of noise. *Journal of Vision*, *6*, 1244–1256.

Carter, O., Konkle, T., Wang, Q., Hayward, V., & Moore, C. (2008). Tactile rivalry demonstrated with an ambiguous apparent-motion quartet. *Current Biology*, *18*, 1050–1054.

Carter, O., & Pettigrew, J. D. (2003). A common oscillator for perceptual rivalries? *Perception*, *32*, 295–305.

Carter, O. L., Presti, D. E., Callistemon, C. Ungerer, Y., Liu, G. B., & Pettigrew, J. D. (2005). Meditation alters perceptual rivalry in Tibetan Buddhist monks. *Current Biology*, *15*, R412–R413.

Chen Y. C., Yeh, S. L., & Spence, C. (2011). Crossmodal constraints on human perceptual awareness: Auditory semantic modulation of binocular rivalry. *Frontiers of Psychology*, *2*, 212.

Chong, S. C., Tadin, D., & Blake, R. (2005). Endogenous attention prolongs dominance durations in binocular rivalry. *Journal of Vision*, *5*(11), 1004–1012. doi:10.1167/5.11.6

Dayan, P. (1998). A hierarchical model of visual rivalry. *Neural Computation*, *10*, 1119–1136.

Donnelly, M., & Miller, R. J. (1995). Ingested ethanol and binocular rivalry. *Investigative Ophthalmology & Visual Science*, *36*, 1548–1554.

Frecska, E., White, K. D., Leonard, C. M., Kuldau, J. M., Bengtson, M., Ricciuti, N., . . . Pettigrew, J. D. (2003). Binocular rivalry in schizophrenia and drug-induced psychosis. *Neuropsychopharmacologia Hungarica*, 5, 4–13.

Frecska, E., White, K. D., & Luna, L. E. (2003). Effects of ayahuasca on binocular rivalry with dichoptic stimulus alternation. *Psychopharmacology*, *173*, 79–87.

Grossberg, S., Yazdanbakhsh, A., Cao, Y., & Swaminathan, G. (2008). How does binocular rivalry emerge from cortical mechanisms of 3-D vision? *Vision Research*, *48*, 2232–2250.

Hohwy, J., Roepstorff, A., & Friston, K. J. (2008). Predictive coding explains binocular rivalry: An epistemological review. *Cognition*, *108*, 687–701.

Jiang, Y., Costello, P., Fangm, F., Huang, M., & He, S. (2006). A gender- and sexual orientation-dependent spatial attentional effect of invisible images. *Proceedings of the National Academy of Sciences USA*, *103*, 17048–17052.

Kang, M. S., Blake, R., & Woodman, G. (2011). Semantic processing does not occur in the absence of awareness induced by interocular suppression. *Journal of Neuroscience*, *31*, 13535–13545.

Kim, C. Y., & Blake, R. (2005). Psychophysical magic: Rendering the visible "invisible." *Trends in Cognitive Sciences*, *9*, 381–388.

Knapen, T., Brascamp, J., Pearson, J., van Ee, R., & Blake, R. (2011). The role of frontal and parietal brain areas in bistable perception. *Journal of Neuroscience, 31*, 10293–10301.

Kovács, I., Papathomas, T., Yang, M., & Feher, A. (1996). When the brain changes its mind: Interocular grouping during binocular rivalry. *Proceedings of the National Academy of Sciences USA*, *93*, 15508–15511.

Lankheet, M. J. M. (2006). Unraveling adaptation and mutual inhibition in perceptual rivalry. *Journal of Vision*, *6*(4), 304–310. doi:10.1167/6.4.1

Lin, Z. C., & He, S. (2009). Seeing the invisible: The scope and limits of unconscious processing in binocular rivalry. *Progress in Neurobiology*, *87*, 195–211.

Lumer, E., Friston, K. J., & Rees, G. (1998). Neural correlates of perceptual rivalry in the human brain. *Science*, *280*, 1930–1934.

Lunghi, C., Binda, P., & Morrone, M. C. (2010). Touch disambiguates rivalrous perception at early stages of visual analysis, *Current Biology*, *20*, R143–R144.

Maruya, K., Watanabe, H., & Watanabe, W. (2008). Adaptation to invisible motion results in low-level but not high-level aftereffects. *Journal of Vision*, *8*(11):7, 1–11. doi:10.1167/8.11.7.

Maruya, K., Yang, E., & Blake, R. (2007). Voluntary action influences visual competition. *Psychological Science*, *18*, 1090–1098.

Meng, M., Remus, D. A., & Tong, F. (2005). Filling-in of visual phantoms in the human brain. *Nature Neuroscience*, *8*, 1248–1254.

Miller, S. M., Hansell, N. K., Ngo, T. T., Liu, G. B., Pettigrew, J. D., Martin, N. G., & Wright, M. J. (2010). Genetic contribution to individual variation in binocular rivalry rate. *Proceedings of the National Academy of Sciences USA*, *107*, 2664–2668.

Montaser-Kouhsari, L., Moradi, F., Zandvakili, A., & Esteky, H. (2004). Orientation-specific adaptation during motion-induced blindness. *Perception*, *33*, 249–254.

Moradi, F., Koch, C., & Shimojo, S. (2005). Face adaptation depends on seeing the face. *Neuron*, *45*, 169–175.

Nguyen, V. A., Freeman, A. W., & Alais, D. (2003). Increasing depth of binocular rivalry suppression along two visual pathways. *Vision Research*, *43*, 2003–2008.

Patterson, R., Winterbottom, M., Pierce, B., & Fox, R. (2007) Binocular rivalry and head-worn displays. *Human Factors, 49*, 1083–1096.

Rao, V. M. (1977). Tilt illusion during binocular rivalry. *Vision Research*, *17*, 327–328.

Sobel, K., & Blake, R. (2003). Subjective contours and binocular rivalry. *Vision Research*, *43*, 1533–1540.

Tong, F. (2005). Investigations of the neural basis of binocular rivalry. In D. Alais & R. Blake (Eds.), *Binocular rivalry* (pp. 63–80). Cambridge, MA: MIT Press.

Tong, F., Meng, M., & Blake, R. (2006). Neural bases of binocular rivalry. *Trends in Cognitive Science*, *10*, 502–511.

Tononi, G., & Edelman, G. M. (2000). Schizophrenia and the mechanisms of conscious integration. *Brain Research Reviews*, *31*, 391–400.

Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience, 8*, 1096–1101.

von Noorden, G. (1990). *Binocular vision and ocular motility: Theory and management of strabismus*. St. Louis, MO: Mosby.

Wade, N. (1980). The influence of colour and contour rivalry on the magnitude of the tilt illusion. *Vision Research*, *20*, 229–233.

Wade, N. (1998). *A natural history of vision*. Cambridge MA: MIT Press.

Wheatstone, C. (1852). Contributions to the physiology of vision—part the second. On some remarkable and hitherto unobserved phenomena of binocular vision. *Philosophical Transactions of the Royal Society*, *142*, 1–17.

Wiesenfelder, H., & Blake, R. (1990). The neural site of binocular rivalry relative to the analysis of motion in the human visual system. *Journal of Neuroscience, 10*, 3880–3888.

Yang, E., Hong, S. W., & Blake, R. (2010). Adaptation aftereffects to facial expressions suppressed from visual awareness. *Journal of Vision*, *10*(12): 24.

Zhou, W., & Chen, D. (2009). Binaral rivalry between the nostrils and in the cortex. *Current Biology*, *19*, 1–5.

Zhou, W., Jiang, Y., He, S., & Chen, D. (2010). Olfaction modulates visual perception in binocular rivalry. *Current Biology*, *20*, 1356–1358.

Chapter 106
Stereo Rotation Standstill and Related Illusions

Max R. Dürsteler and Erika N. Lorincz

DESCRIPTION OF THE ILLUSION

The Interactive Figure VIII.106-1 on the Companion Website encodes depth information in a red-cyan stereogram requiring the use of anaglyphic eye glasses (red filter before the left eye). When the display shows noticeable stutter, one can reduce the window's size to decrease the computing load. The application shows a wheel-like three-dimensional (3D) structure made out of eight convex sectors. The cross-section of each sector (i.e., depth) is an approximation of the positive half-period of a sine function (Fig. VIII.106-1).

The 3D wheel ("3D structure") is rotating around its axis perpendicular to the computer's screen. The random-dot texture used to encode the wheel's depth (Julesz, 1960) is static; that is, the same texture is used in each frame. There are, however, some small, horizontal movements of the texture's elements due to disparity changes during the wheel's rotation. The application allows choosing between different 3D structures (VIII.106-2a–f) and different textures (Fig. VIII.106-2g–j). Readers may want to adjust parameters such as the disparity of the wheel ("maximal disparity" in picture elements ["pels"]).

When looking at the center of the rotating wheel made out of eight sectors, the overall impression is that it is standing *still*. This is called the stereo-rotation standstill illusion (SRSI). Even when we increase the velocity to 60°/s, the wheel still appears motionless, yet the details of the 3D structure (sectors) can clearly be distinguished. When tracking a sector near the wheel's rim, one perceives that it is actually moving and may therefore infer the 3D wheel's rotation.

SIGNIFICANCE OF THE ILLUSION

Standstill illusions have been found to occur for translational motion in colored vision (Lu, Lesmes, & Sperling, 1999; Ramachandran & Gregory, 1978) and in stereo vision (Julesz & Payne, 1968; Tseng, Gobell, Lu, & Sperling, 2006). Translational stereo-motion standstill illusions were produced using dynamic random-dot stereograms of a rapidly moving 3D grating. The grating was perceived as completely motionless, only when its velocity was above medium to high thresholds (e.g. 4 to 7 Hz in Tseng et al., 2006), while its structural details were still discernible.

However, our SRSI seems different from translational stereo-motion standstill. First, the motionless percept is obtained with much lower temporal frequencies in our stereo-rotation display than in translational motions (e.g., for a wheel with eight spokes rotating at 12°/s the temporal frequency is 0.267 Hz). Second, while we do not perceive the whole wheel's rotation at such low angular velocities, we still clearly perceive local stereo motion (e.g., when tracking one of the rotating wheel's sectors). Obviously the ability to perceive local stereo motion cannot guarantee the perception of complex stereo motions such as stereo rotation.

The robustness of the SRSI strongly suggests that the detectors of local stereo motion in our visual system do not feed into detectors of more global general rotation detectors and/or that specialized detectors of stereo rotation are very scarce or absent. This latter hypothesis fulfills the criteria of "too weak

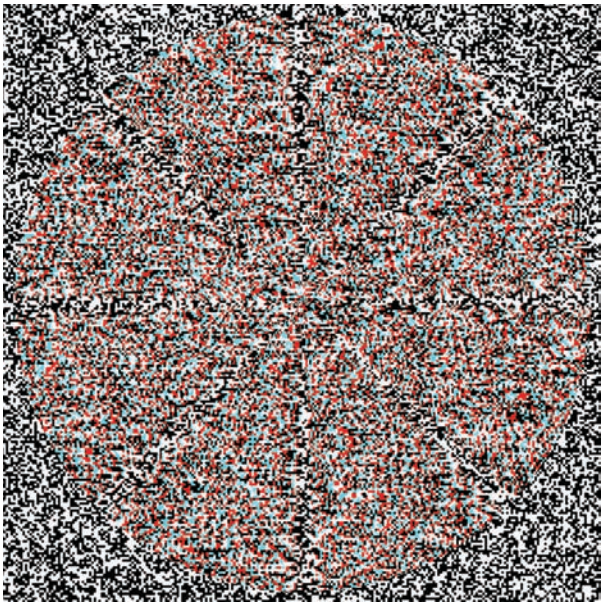
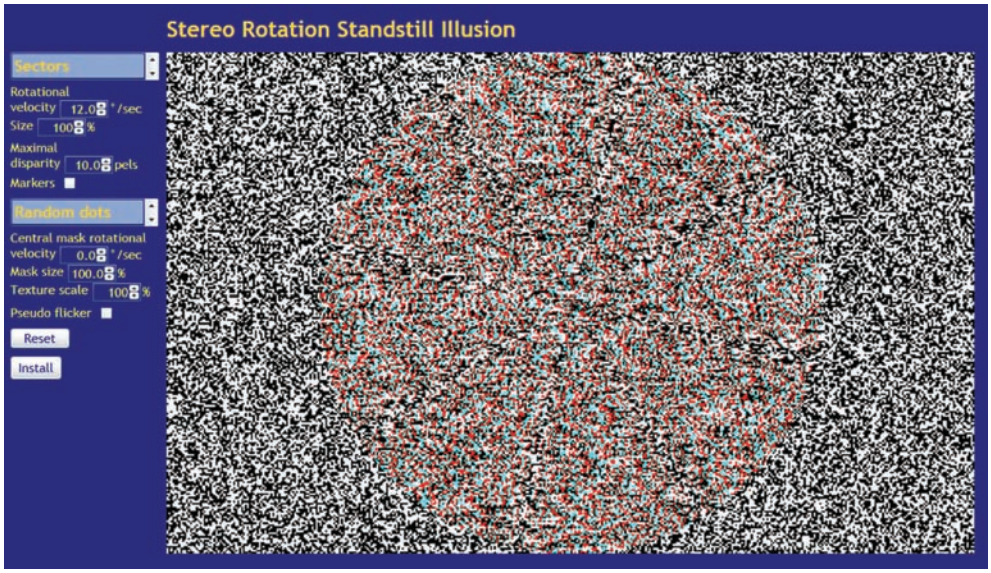


Figure VIII.106-1. Red-cyan anaglyph of a wheel with eight sectors. We used red-cyan anaglyph eyeglasses to perceive the hidden 3D wheel. With the red filter before the left eye and the cyan filter before the right eye, the sectors of the wheel are perceived as popping out of the figure's plane, while with the reversed filters' arrangement, the wheel's sectors are perceived as embossed. (Figure by Max R. Dürsteler.)



Interactive Figure VIII.106-1. Figure by Max R. Dürsteler.

motion signal for the motion system to be resolved", proposed by Lu et al. (1999) in their general definition of motion standstill. In that respect, the SRSI belongs to "motion standstill" illusions, despite the fact that it is observed for lower velocities than those suggested by these authors.

However, we do not know whether specialized detectors of stereo rotation exist. In addition, first-order motion inputs might participate in object–ground (or other object) segregation, thus providing some cues to rotational motion, so that parts of the scene should appear as rotating and others as stationary. In this chapter we try to explain how the visual system deals with complex stereo-motions illusions such as the SRSI.

VARIANTS AND RELEVANT PARAMETERS

The SRSI and its variants can be elicited using Interactive Figure VIII.106-1. The settings to obtain the percepts described below are listed in Table VIII.106-1 as conditions (a) to (o), except for the size of the 3D wheel (use 100%), the maximal disparity (10 pels), and the texture scale (100%).

Stereo-Rotation Illusions With Static Texture

SRSI

The SRSI can be best seen when fixating the center. To convince yourself that the 3D structure is really rotating, check the "markers" box in Interactive Figure VIII.106-1. Eight yellow markers appear underneath the semitransparent sectors of the wheel that actually rotate with it. The wheel still appears as standing still.

Typical ("Physical") Stereo Rotation

When the luminance-defined (random-dot) texture starts rotating together with the 3D wheel it encodes, one immediately obtains a clear percept of a turning wheel.

Stereo-Rotation Capture

When the overlying texture rotates while the 3D wheel is motionless, the wheel appears to corotate with its surface's texture. The motion of the overlying texture captures the motion of the 3D wheel regardless of its physical rotation (i.e., its velocity and rotation direction; e.g., $\omega_w = -12°/s$ and $\omega_T = +12°/s$). However, on close inspection of an intersectorial edge near the wheel's periphery, we see that it is actually the dots' texture that is rotating underneath the wheel's sector. This allows inference of the existing differences between the angular velocities of the texture and that of the wheel.

Partial Stereo-Rotation Capture

When only a portion of the overlaying texture rotates with the 3D wheel, one perceives a kind of surrealistic display with a frozen outer ring and an inner turning disk, where one portion of the very same sectors appears to rotate, even if both parts never split. Translating the axis of the texture's rotation away from the axis of wheel's rotation ("Mask's center of rotation": horizontal offset or vertical offset) shows that the location of the texture's rotational axis determines the location of the perceived common axis of rotation (use random-dot texture and $\omega_T > 1°/s$).

Stereo-Rotation Illusions With Flickering Texture

The static random-dot texture used to encode the rotating 3D wheel seemingly overrides any input from a presumed stereo-rotation detector, giving rise to the percepts of SRSI and stereo rotation capture illusion. Consequently, we tried to implement the recommended method of Julesz and Payne (1968), which normally allows higher-order stereo motion to be perceived in the absence of first-order luminance signals in the texture, to check whether our illusions would disappear. To avoid any interference with the depth-encoding texture, they advise to use *dynamic random-dot stereograms*. Such a *flickering* stereo display does, by

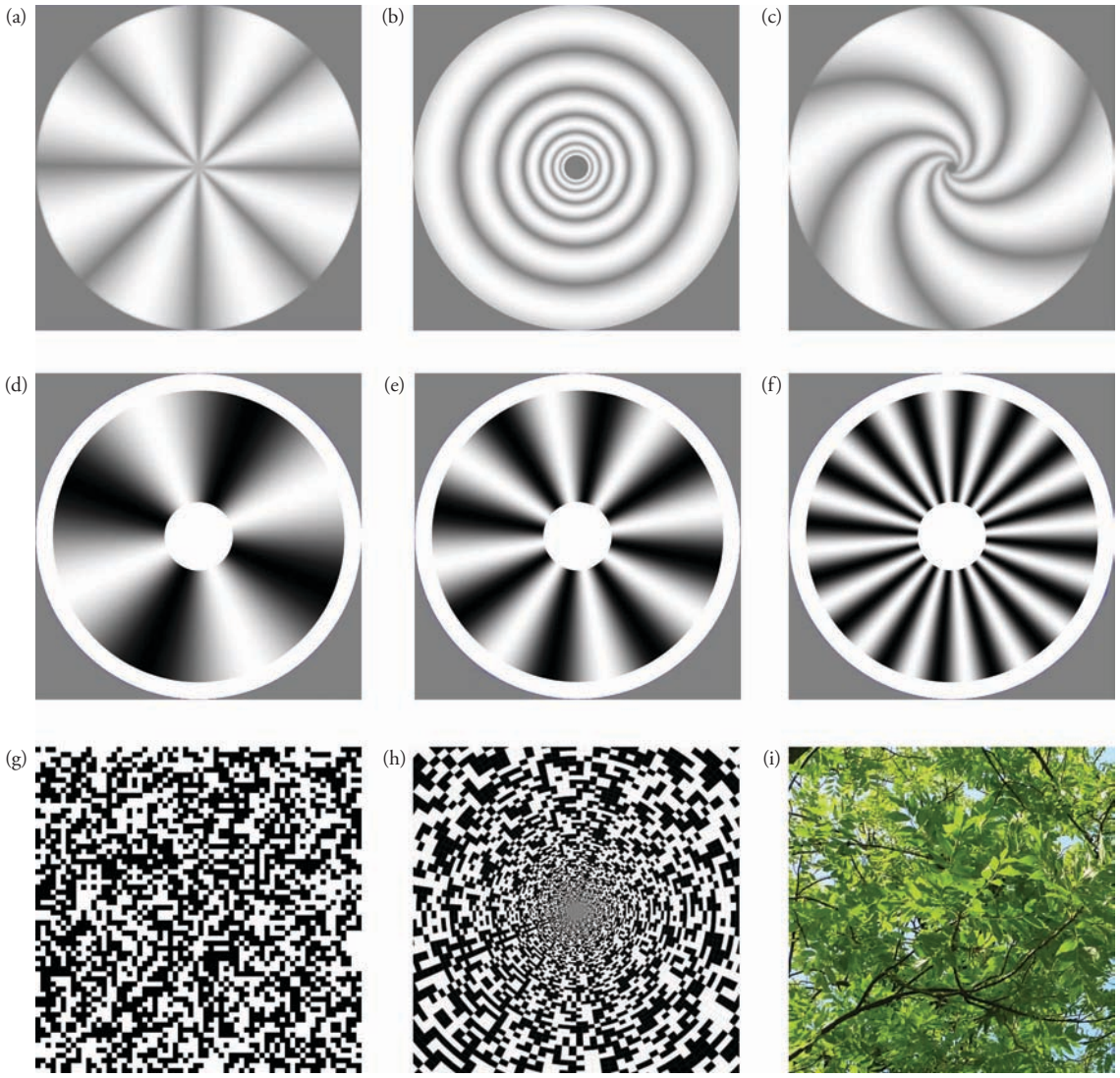


Figure VIII.106-2. Depth maps and textures used by the accompanying application. a, b, c: Depth map of wheels made out of eight sectors (a), concentric rings (b), or six logarithmic right-handed spirals (c). One sector is made of 32 subsectors whose depths are varied according to the positive half-period of a sinusoid. Concentric rings' diameter increases exponentially in size. d, e, f: Depth maps of wheels with 4, 8, or 16 spokes, respectively. g, h, i: Random dots (g) periodic random dots (h), and photograph of an ash tree (i) used as textures. For random-dot textures of the B. Julesz type, each texture's element has an equal chance of being black or white. The periodic random-dot texture is made of rings whose diameters increase exponentially in size. Each ring is covered with the same random-dot texture (size 4 • 240 texture's elements). (Figure by Max R. Dürsteler.)

definition, not rotate, since the dots in successive frames are uncorrelated; the only correlation that remains is the one between dots seen by the left and the right eyes. Due to computing time and memory limitations, our application only "imitates" flicker by irregularly shifting and rotating the same random-dot texture on each frame.

SRSI Flicker

Despite the flickering texture, the wheel appears to be standing still. Again, tracking one of the wheel's part helps to build up a percept of a somewhat jerky rotation.

Stereo-Rotation Capture by Markers

Compare the static percept we have seen already with the nice, smooth rotation of the wheel with the yellow markers. Apparently some help from a luminance-based first-order system is enough to reach the threshold to perceive complex stereo motion with a dynamic random-dot display.

Stereo-Scaling Motion Illusions

By "scaling" we mean the actual expansion or contraction. To produce geometrically accurate 3D scaling and/or infinite texture motion, we used periodically repeated concentric circular rings for the 3D wheel (Fig. VIII.106-2b) and the random-dot texture (Fig. VIII.106-2h). All the dimensions of the rings, such as the radius, the thickness, and the depth's profile, are increased from the innermost to the outermost ring according to a geometric sequence, which goes from the inner smallest to the outermost largest value. The nth term of a geometric series with an initial value a and a common ratio r is given by the formula: $a_n = ar^{n-1}$.

Table VIII.106-1. Critical parameters for complex stereo-motion illusions. w = 3D wheel; t = texture; Ω = angular velocity; F = temporal frequency of the expanding rings or of the periodic random dots.

	$Type_w$	Ω_w (°/Sec)	F_w (Hz)	*Markers*	$Type_t$	Ω_t (°/Sec)	F_t (Hz)	$Size_t$ (%)	*Flicker*	*Percepts*
a	Sectors	12	n.a.	no/yes	Ran. Dots	0	n.a.	100	no	SRSI
b	Sectors	12	n.a.	no	Ran. Dots	12	n.a.	100	no	Stereo rotation
c	Sectors	0	n.a.	no/yes	Ran. Dots	12	n.a.	100	no	SRCI
d	Sectors	12	n.a.	no	Ran. Dots	12	n.a.	70	no	Partial stereo-rotation capture
e	Sectors	12	n.a.	no	Ran. Dots	0	n.a.	100	yes	SRSI with flicker
f	Sectors	12	n.a.	yes	Ran. Dots	0	n.a.	100	yes	Stereo-rotation capture by markers
g	Rings	n.a.	0.4	no/yes	Ran. Dots	0	n.a.	100	no/yes	Stereo-contraction standstill
h	Rings	n.a.	–0.4	no/yes	Ran. Dots	0	n.a.	100	no/yes	Stereo-expansion standstill
i	Rings	n.a.	0.4	no	Periodic RD	n.a.	0.4	100	no	Stereo-contraction
j	Rings	n.a.	–0.4	no	Periodic RD	n.a.	–0.4	100	no	Stereo-expansion
k	Rings	n.a.	0	no	Periodic RD	n.a.	0.4	100	no	Stereo-contraction capture
l	Rings	n.a.	0	no	Periodic RD	n.a.	–0.4	100	no	Stereo-expansion capture
m	Spirals	12	n.a.	no/yes	Ran. Dots	0	0	100	no/yes	Stereo-spiraling motion standstill
n	Spirals	12	n.a.	no	Periodic RD	12	0.4	100	no	Stereo-spiraling motion
o	Spirals	0	n.a.	no	Periodic RD	12	0.4	100	no	Stereo-spiraling motion capture

During expansion, rings and/or circular periodical random-dot patterns are scaled up in size by increasing the scaling factor from a to ar or to ar^k, where k is a positive integer. Then the initial rings' display is restored, and the rings expand again. The outermost and the innermost rings are masked off to reach the percept of a continuous expansion. A similar scheme is used for contraction. The center has been masked off, as the rings become infinitely small to avoid moiré artifacts.

Stereo-Contraction Standstill

Despite that it takes 2.5 s for rings to travel from the initial to the second outermost position when they contract at 0.40 Hz, the rings appear not to contract at all when using the static random-dot display ("pseudo flicker" unchecked). However, when we track one of the rings we are able to infer their continuous contraction. Checking the "markers" box allows yellow dashed lines perpendicular to the 3D rings to travel physically with the rings' ridges. Despite these luminance cues, the rings appear either as stationary or to faintly contract. The contracting "perpendicular" yellow markers partially fail to capture the stereo rings. If we use instead luminance markers "parallel" to the contracting rings (check "circles"), the 3D rings appear to contract again together with the luminance circles.

If we have the "pseudo flicker" and remove the luminance cues ("markers" or "circles"), the contractions of the 3D rings again disappear. If we add again the luminance cues, the same percept as with the static texture reappears (i.e., a weak feeling of contraction of the 3D structure with the perpendicular markers and a clear contraction with the parallel luminance circles). In other words, only the yellow circles appear to capture the motion of the stereo rings.

Stereo-Expansion Standstill

The same phenomena described previously also occur with expansion.

Typical Stereo Contraction

When both the periodical random-dot texture and the 3D rings are contracting together, one perceives a smooth contraction of the rings with the texture, together with a motion in depth component directed away from the observer.

Typical Stereo Expansion

A comparable phenomenon can be observed with expansion accompanied by a motion in depth component directed toward the observer.

Stereo-Contraction Capture

The percept of contracting rings remains, despite that they are in fact stationary. Local inspection of a ring's ridge reveals that the fixated ring is stationary but the other rings are still perceived as contracting. Luminance cues provided by either perpendicular "markers" or parallel "circles" do not help much to perceive the 3D rings as stationary.

Stereo-Expansion Capture

The 3D rings appear as expanding despite the fact that they are stationary.

Stereo-Spiral Motion

Our application uses logarithmic spirals (Fig. VIII.106-2c). In polar coordinates (r is the radius, and θ is the angle of a point in the spiral), the formula for logarithmic right-handed spiral can be written as $\theta = \frac{1}{b} \ln\left(r/a\right)$, with a and b being arbitrary positive real constants. In parametric form, the formulas of the logarithmic spiral are $x(t) = ae^{bt}\cos(t)$ and $y(t) = ae^{bt}\sin(t)$, with t being an arbitrary rational value. To obtain a left-handed spiral, set the constant b to a negative value. In our demonstration, the size of a forever growing spiral is limited by a circular clipping region. Note that logarithmic spirals have some uncommon features such that scaling transformations also result in spirals' rotations and rotation's transformation in spirals' expansion or contraction.

Stereo-Spiraling Motion Standstill

The physically rotating 3D spirals are perceived as stationary, even with the yellow markers turned on. The flickering texture (check "pseudo flicker") provokes exactly the same percept as for rotation and scaling (i.e., no clear percept of rotation without luminance markers) and a percept of smooth rotation with yellow markers.

Typical Stereo-Spiraling Motion

When texture and spirals are in synchrony, one perceives a smooth spiraling motion rotating clockwise and contracting together with a motion in the depth component going away from the observer.

Stereo-Spiraling Motion Capture

When only the texture is moving, the physically stationary 3D spirals are perceived as rotating and contracting.

DISCUSSION

The SRSI and related complex stereo-motion standstill illusions such as scaling and spiraling occur when we fixate the center of a physically moving 3D structure whose stereo cues are encoded with a static random-dot texture. Even using dynamic random-dot flicker, which normally allows higher-order (stereo) motion to be perceived instead of the static texture, does not help to discriminate the real motion of the 3D structure. Only adding "parallel" luminance cues moving together with the 3D structure allows the real motion to be perceived. However, desynchronizing the motion of these markers (e.g., stopped or opposite motion) from that of the 3D structure reveals the capturing effect of luminance on the structure's motion.

Yet luminance capture can be overridden by texture motion. Using a static (as opposed to a dynamic) but coherently moving random-dot texture (periodic or not) generates a percept of congruent motion of the physically static (or oppositely moving) 3D structure together with the texture, and this even if the luminance cues are present. In the latter case, they appear as moving on their own.

We can draw three assumptions from the reported observations. First, the visual percept is governed according to a hierarchy in the systems in which the luminance variations are coded (i.e., either markers ["blobs"], textures, or 3D structure). Following decreasing levels of impact, coherent motion of the texture overrules any other luminance motions, whether it determines markers or 3D structures' motions. In turn, motion of luminance markers overrides any variation of luminance coming from either dynamic (flickering) texture or 3D moving structure. Last, static variation of texture luminance prevails over 3D-structure's motion to produce the visual percept. Second, following this idea, motion standstill and motion capture originate both from a same phenomenon, that is, a capturing effect of the texture, static for the former and moving for the later. Finally, it is likely that these illusions are due to scarcity or even complete absence of complex stereo-motion detectors.

Complex Stereo-Motion Standstill

We found three new examples of motion standstill as defined by Lu et al. (1999) for color vision and Tseng et al. (2006) for stereo vision. The motion standstill illusions we observed were not due to a complete failure to perceive any stereo motion: the fixation of a single element of the 3D wheel allowed perception of local stereo motion. However, global complex stereo motions were impossible to recognize. While "simple motion" refers to translational or unidirectional motion (i.e., one-dimensional), "complex motion" refers to motions in two dimensions, such as rotation, scaling (i.e., expansion or contraction), spiraling, and shearing (Beardsley & Vaina, 2005; Lee & Lu, 2010; Morrone, Burr, & Vaina, 1995).

The failure to perceive some types of motions does not demonstrate, however, the inability of the visual system to treat these signals (Patterson, 1999). It could be due to the presence of an inhibiting element in the stimulus used. For instance, it has been shown that the presence of high-contrast random-dot pattern suppresses the perception of motion in a low-contrast or equiluminant (peripheral) target (Murakami, 1999; Murakami & Shimojo, 1993). The effect is known as "position capture." However, since a percept of smooth complex stereo motion is only achieved with a little help from the luminance system, we have doubts about the existence of independent global rotational, scaling, or spiral detectors for stereo vision.

Complex Stereo-Motion Luminance Capture

Motion-capture phenomena, whereby a dominant high-contrast luminance stimulus (e.g., our markers) captures an equiluminant target, have already been described for simple (Culham & Cavanagh, 1994; Murakami, 1999; Murakami & Shimojo, 1993) and complex motions (Duijnhouwer, van Wezel, & van den Berg, 2008; Lee & Lu, 2010). We observed motion-capture phenomena of stereo targets with either luminance blobs (the markers) or luminance textures. Stereo motion captured by luminance blobs is best seen with long-inducing markers oriented parallel to the 3D-structure ridges encoded by a dynamic (flickering) rather than by a static random-dot texture. In contrast to luminance markers, which can capture the whole 3D structure, coherent motion of a luminance texture captures only the motion of the underlying parts and not the entire 3D structure.

However, luminance texture motion abolishes any motion-capture effects of moving luminance markers. Why are the latter less successful in capturing motion than the former? A possible explanation would be that luminance markers have borders defined in a unique visual modality (i.e., luminance differences) and are therefore perceived as independent objects moving on their own. By contrast, the apparent borders of our moving luminance textures are defined in two different visual modalities (i.e., disparity jumps and differential motion). Consequently, the combination of the moving texture with the underlying 3D structure would be perceived as a unique (indivisible) object animated by a common motion.

Complex First-Order Motion and Complex Stereo Motion

The fact that in our examples of complex stereo-motion standstill local stereo motion can still be perceived tends to demonstrate that different motion systems must be involved in the perception of "local" translational stereo motion and in "global" complex stereo motions, such as rotation, spiral, or scaling. Motion stimuli can be classified according to their complexity. One commonly used classification is based on the order of statistics required to detect stimulus motion (Cavanagh & Mather, 1989). First-order statistics define the frequency with which specific combinations of intensity values occur at *one point* (e.g., changes in luminance). Second-order statistics define the frequency with which specific combinations of intensity values occur at a *pair of points* (e.g., changes in stereo depth that can only be detected by comparing inputs from the left and the right eyes). While stereo motion is a second-order motion (Patterson, 1999; Seiffert & Cavanagh, 1998), it has been argued that our visual system uses a common third-order system, which can detect motion of salient features in stereo and/or equiluminant color stimuli (Lu & Sperling, 1995, 2001; Tseng et al., 2006).

There is convincing psychophysical evidence that complex first-order (luminance-based) motion stimuli are processed in two steps (Burr, Concetta Morrone, & Vaina, 1998; Morrone et al., 1995). The outputs of small field receptors such as those known to exist in V1 up to the mediotemporal (MT) area are integrated by a specialized large receptive field motion mechanism known to operate in extrastriate cortical areas upstream to MT. Coherent with this idea, neurons have been found in the dorsal region of the medial superior temporal area (MSTd) of the monkey's cerebral cortex that respond to radial, circular, and spiral motions (Duffy & Wurtz, 1991a, 1991b; Graziano, Andersen, & Snowden, 1994; Saito et al., 1986; Tanaka, Fukada, & Saito, 1989; Tanaka & Saito, 1989). In addition, about 90% of MSTd neurons are disparity sensitive (Roy & Wurtz, 1990; Roy, Komatsu, & Wurtz, 1992). In humans, full-field optic-flow stimuli have been shown to activate the corresponding region of the MST+ complex (Dukelow et al., 2001). No study, however, has so far evidenced the existence of detectors that combine complex motion and disparity processing.

Consistent with the idea of scarcity or inexistence of complex stereo-motion detectors, response thresholds to complex second-order contrast-defined stimuli are found to be much higher than for first-order stimuli (Allen & Derrington, 2000; Bertone & Faubert, 2003). In addition, we did not find any clear-cut evidence for the existence of stereo-motion inputs to the complex motion system, as we never observed smooth complex motion of the 3D structure without the capturing effect of luminance-defined texture or markers, even when using our dynamic random-dot stereograms. Seemingly even the hypothetical common third-order motion system based on saliency (Lu & Sperling, 2001) could not detect the complex motion of our stereo stimuli rich in salient features. If the detection of complex motions by our visual system relies exclusively on inputs from the first-order system, it is likely to give rise to illusions such as "complex stereo-motion standstill" and "complex stereo-motion capture" whenever we look at a scene with moving objects containing an incongruence between the complex motion of luminance (texture or markers) and the stereoscopic (3D-structure) motion. However, such a situation is very unlikely to occur in daily life (although watch, e.g, leaves of lilies in a pound while an expanding water wave passes by).

To conclude, the complex stereo-motion standstill and the complex stereo-motion capture illusions help to shed light on how the visual system combines luminance and disparity information from a rotating, contracting/expanding, and/or spiraling 3D-structure in order to reach a unified percept of the object's motion.

REFERENCES

Allen, H. A., & Derrington, A. M. (2000). Slow discrimination of contrast-defined expansion patterns. *Vision Research*, *40*(7), 735–744. doi:10.1016/s0042-6989(99)00223-0

Beardsley, S. A., & Vaina, L. M. (2005). Psychophysical evidence for a radial motion bias in complex motion discrimination. *Vision Research*, *45*(12), 1569–1586. doi:10.1016/j.visres.2004.11.025

Bertone, A., & Faubert, J. (2003). How is complex second-order motion processed? *Vision Research*, *43*(25), 2591–2601. doi:10.1016/s0042-6989(03)00465-6

Burr, D. C., Concetta Morrone, M., & Vaina, L. M. (1998). Large receptive fields for optic flow detection in humans. *Vision Research*, *38*(12), 1731–1743. doi:10.1016/s0042-6989(97)00346-5

Cavanagh, P., & Mather, G. (1989). Motion: The long and short of it. *Spatial Vision*, *4*(2–3), 103–129.

Culham, J. C., & Cavanagh, P. (1994). Motion capture of luminance stimuli by equiluminous color gratings and by attentive tracking. *Vision Research*, *34*(20), 2701–2706.

Duffy, C. J., & Wurtz, R. H. (1991a). Sensitivity of MST neurons to optic flow stimuli: I. A continuum of response selectivity to large-field stimuli. *Journal of Neurophysiology*, *65*(6), 1329–1345.

Duffy, C. J., & Wurtz, R. H. (1991b). Sensitivity of MST neurons to optic flow stimuli: II. Mechanisms of response selectivity revealed by small-field stimuli. *Journal of Neurophysiology*, *65*(6), 1346–1359.

Duijnhouwer, J., van Wezel, R. J. A., & van den Berg, A. V. (2008). The role of motion capture in an illusory transformation of optic flow fields. *Journal of Vision*, *8*(4): 27, 21–18.

Dukelow, S. P., DeSouza, J. F., Culham, J. C., van den Berg, A. V., Menon, R. S., & Vilis, T. (2001). Distinguishing subregions of the human MT+ complex using visual fields and pursuit eye movements. *Journal of Neurophysiology*, *86*(4), 1991–2000.

Graziano, M., Andersen, R., & Snowden, R. (1994). Tuning of MST neurons to spiral motions. *The Journal of Neuroscience*, *14*(1), 54–67.

Julesz, B. (1960). Binocular depth perception of computer-generated patterns. *Bell System Technical Journal*, *39*, 1125–1162.

Julesz, B., & Payne, R. A. (1968). Differences between monocular and binocular stroboscopic movement perception. *Vision Research*, *8*(4), 433–444.

Lee, A. L. F., & Lu, H. (2010). A comparison of global motion perception using a multiple-aperture stimulus. *Journal of Vision*, *10*(4): 9, 1–16.

Lu, Z.-L., Lesmes, L. A., & Sperling, G. (1999). Perceptual motion standstill in rapidly moving chromatic displays. *Proceedings of the National Academy of Scences USA*, *96*(26), 15374–15379.

Lu, Z.-L., & Sperling, G. (1995). The functional architecture of human visual motion perception. *Vision Research*, *35*(19), 2697–2722. doi:10.1016/0042-6989(95)00025-u

Lu, Z.-L., & Sperling, G. (2001). Three-systems theory of human visual motion perception: Review and update. *Journal of the Optical Society of America A 18*(9), 2331–2370.

Morrone, M. C., Burr, D. C., & Vaina, L. M. (1995). Two stages of visual processing for radial and circular motion. *Nature*, *376*(6540), 507–509.

Murakami, I. (1999). Motion-transparent inducers have different effects on induced motion and motion capture. *Vision Research*, *39*(9), 1671–1681. doi:10.1016/s0042-6989(98)00195-3

Murakami, I., & Shimojo, S. (1993). Motion capture changes to induced motion at higher luminance contrasts, smaller eccentricities, and larger inducer sizes. *Vision Research*, *33*(15), 2091–2107.

Patterson, R. (1999). Stereoscopic (cyclopean) motion sensing. *Vision Research*, *39*(20), 3329–3345. doi: 10.1016/s0042-6989(99)00047-4

Ramachandran, V. S., & Gregory, R. L. (1978). Does colour provide an input to human motion perception? *Nature*, *275*(5675), 55–56.

Roy, J., Komatsu, H., & Wurtz, R. (1992). Disparity sensitivity of neurons in monkey extrastriate area MST. *The Journal of Neuroscience*, *12*(7), 2478–2492.

Roy, J.-P., & Wurtz, R. H. (1990). The role of dispaity-sensitive cortical neurons in signaling the direction of self-motion. *Nature*, *348*, 160–162.

Saito, H., Yukie, M., Tanaka, K., Hikosaka, K., Fukada, Y., & Iwai, E. (1986). Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *The Journal of Neuroscience*, *6*(1), 145–157.

Seiffert, A. E., & Cavanagh, P. (1998). Position displacement, not velocity, is the cue to motion detection of second-order stimuli. *Vision Research*, *38*(22), 3569–3582. doi: 10.1016/s0042-6989(98)00035-2

Tanaka, K., Fukada, Y., & Saito, H. A. (1989). Underlying mechanisms of the response specificity of expansion/contraction and rotation cells in the dorsal part of the medial superior temporal area of the macaque monkey. *Journal of Neurophysiology*, *62*(3), 642–656.

Tanaka, K., & Saito, H. (1989). Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *Journal of Neurophysiology*, *62*(3), 626–641.

Tseng, C. H., Gobell, J. L., Lu, Z. L., & Sperling, G. (2006). When motion appears stopped: Stereo motion standstill. *Proceedings of the National Academy of Sciences USA*, *103*(40), 14953–14958.

Chapter 107
The Graph-Paper Effect

A Moving, Illusory, Stereoscopic Texture

Mark Georgeson

INTRODUCTION

Most visual illusions involve distorted or altered perception of objects or events, or misinterpretation of image information. The discrepancy between what we experience and what is physically present in the world or in the retinal image can be large, surprising, and dramatic. It is much rarer to see things that simply are not there at all. The experience of hallucinations is often linked to neural effects of drugs or alcohol, psychiatric conditions, or brain pathology, but several more ordinary conditions can also make us see things that are not there. Dreams are one example, but in waking life repetitive stimuli, such as grating patterns or flickering lights, can induce perception of a range of illusory geometric patterns, forms, and movements—during or after exposure to the inducing stimulus. This chapter describes one such illusory phenomenon—the *graph-paper effect*—that sheds light on "low-level" or "mid-level" visual processes, specifically the inhibitory interactions between neurons that encode contours, orientation, motion, and stereo depth. It is striking and surprising in its own right (Georgeson, 1980a) and is related to a range of other visual effects and aftereffects (Georgeson, 1976a, 1976b, 1980b; MacKay, 1957).

HOW TO EXPERIENCE THE GRAPH-PAPER EFFECT

The clue is in the name. Many forms of *plaid* formed from two superimposed grid patterns (*gratings*) will do. So will some checked shirts. Graph paper is a form of plaid but is less readily available than it used to be, so Figures VIII.107-1 to VIII.107-4 show several plaids that were synthesized by computer in an effort to optimize the effect, based on previous experimental data (Georgeson, 1980a). The PDF file of these figures (Interactive Figure VIII.107-1) is best viewed on a computer screen using a PDF viewer such as Preview or Adobe Reader that allows image size and orientation to be easily varied. Expand the image window to fill the screen, to center the image, and remove background clutter. The effect is somewhat fragile, and this lack of competing clutter seems to help.

Use the zoom buttons (+/–) to set the image size so that the lines are spaced about 0.5 to 1 cm apart and view the image from about 40 to 50 cm. Our goal now is to get your point of gaze to track smoothly, horizontally, back and forth *in front of* the screen image. Hold a pencil or pen vertically about 2 to 3 cm in front of the screen, so that its tip is roughly central in the image. Move the tip back and forth horizontally, parallel to the screen, maintaining a constant distance 2 to 3 cm in front of the screen. Speed of movement should be about 2 to 3 cm/s. Keep your head steady and follow the pen tip with your eyes.

With luck, and some practice, you will see the graph-paper effect. Your visual field should be awash with faint, illusory lines that move back and forth with your gaze. Experiment with different speeds of tracking movement and image size to try to maximize the effect. It may help to pay attention to the illusory lines and ignore the image on screen. Use the up and down arrow keys to try the same

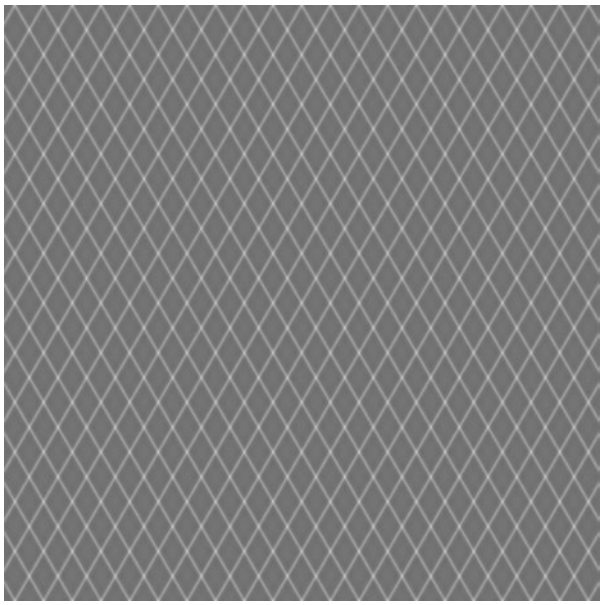
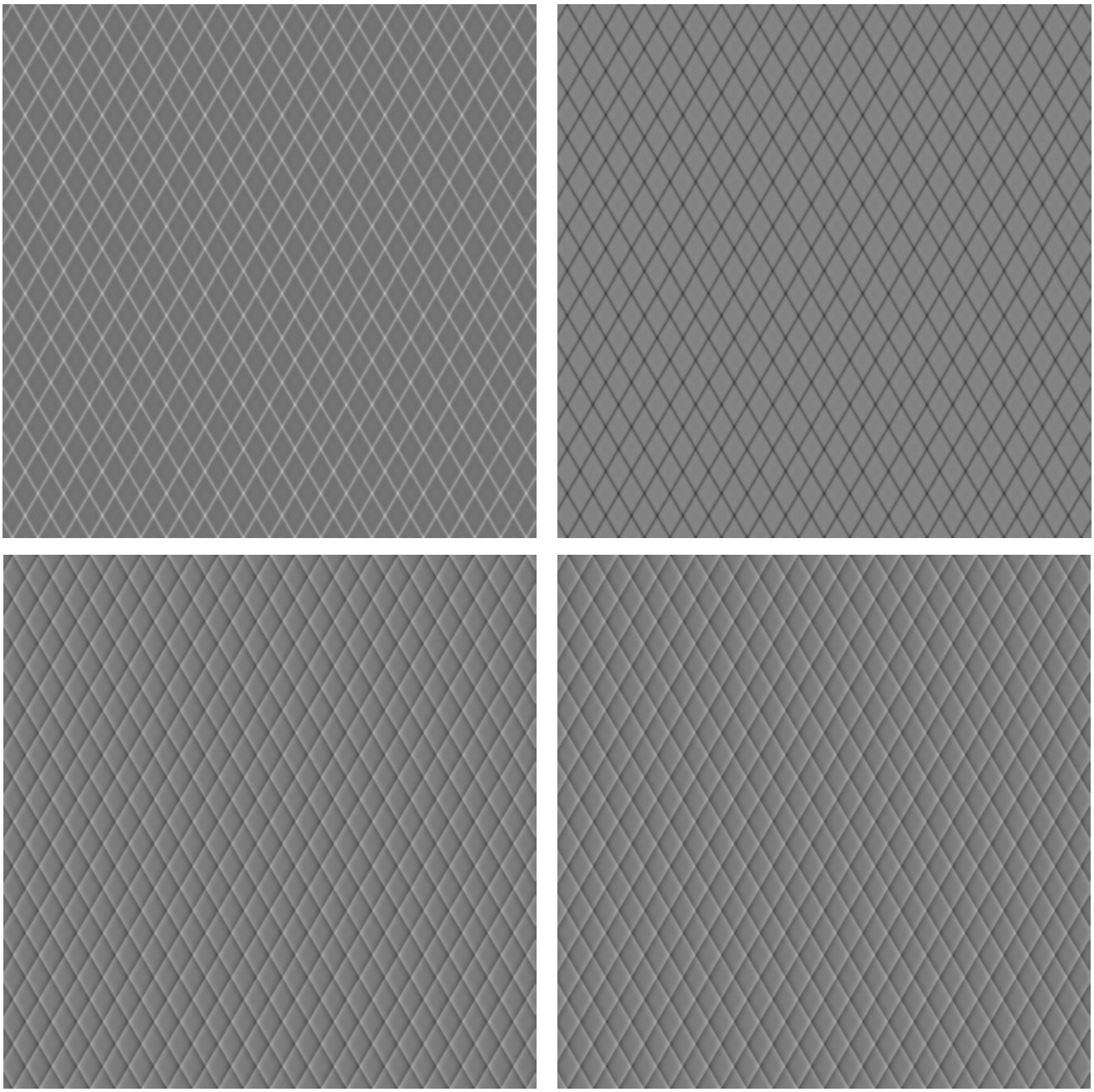


Figure VIII.107-1. Light lines, medium contrast, ±30° from vertical. (Figure by Mark Georgeson, 2015.)



Interactive Figure VIII.107-1. Figure by Mark Georgeson, 2015.

procedure with Figures VIII.107-2, VIII.107-3, and VIII.107-4. If you see the effect but then seem to lose it with repeated viewing, take short breaks between viewings.

CHARACTERISTICS OF THE GRAPH-PAPER EFFECT

Motion

Movement of the inducing image is necessary. If you stop moving your gaze, the illusory lines largely disappear. They move smoothly with your gaze but never spill outside the area of the plaid that induces them. Their speed matches your tracking speed, suggesting that the lines are effectively stationary in retinal coordinates. In this respect, their perceived motion is like that of the conventional negative afterimage, which also appears to move with the eyes. Both are stationary with respect to the retina, but they are moving relative to the image of the stationary world that moves across your retina when your eyes move.

If the on-screen image itself is moved, with the eyes fixing on a stationary point, the effect can still be seen. This shows that retinal movement of the inducing image is the key factor, rather than eye movement per se.

Orientation

In the conditions described so far, the dominant orientation of the illusory lines should appear to be vertical, with

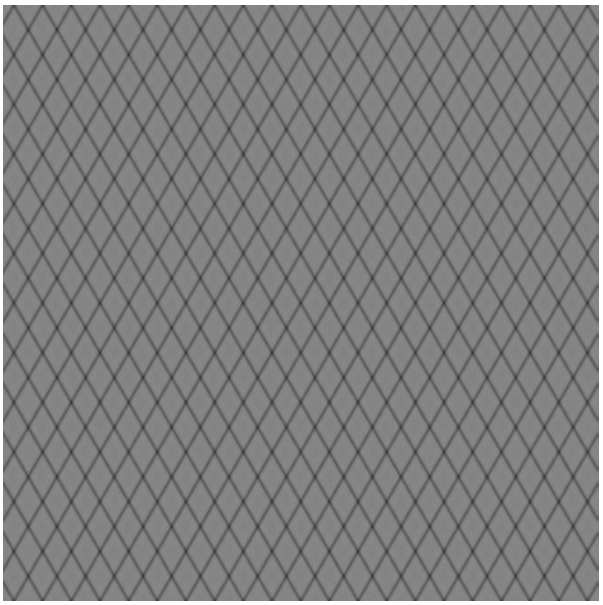


Figure VIII.107-2. Dark lines, medium contrast, ±30° from vertical. (Figure by Mark Georgeson, 2015.)

horizontal movement. Close inspection, however, may reveal other sets of lines oriented about ±60° to ±70° from vertical, forming a noisy plaid that also moves horizontally. Try rotating the image through 90° in your PDF viewer, and then tracking vertically instead of horizontally. The illusory lines should now be predominantly horizontal, and moving vertically.

Interactive Figure VIII.107-2 is a multipage PDF that you can view the same way using the arrow keys to step through different plaid angles (marked in the top left-hand corner of each image). All the images are light lines; only the angle between them varies. You should find that the graph-paper effect is most prominent at intermediate angles from ±30° to ±45° (Georgeson, 1980a).

Edge Polarity

I have so far called them illusory *lines*, but close inspection should reveal that they often look like *edges,* with a specific polarity: light-to-dark or dark-to-light when reading from left to right. One might guess that there is some opposition here between two classes of visual feature—lines and edges. Perhaps inducing lines create illusory edges and vice versa; however, this seems not to be so, as Figures VIII.107-3 and VIII.107-4 demonstrate. Instead of light or dark lines (Figs. VIII.107-1 and VIII.107-2), Figure VIII.107-3 contains contours that we might call "edgelets," with dark-to-light polarity everywhere, while Figure VIII.107-4 has edgelets with light-to-dark polarity. All four images are in Interactive Figure VIII.107-1. Figure VIII.107-5 illustrates the luminance profiles of the four stimulus features to make these distinctions clearer. Figures VIII.107-3 and VIII.107-4 should confirm that edgelets of one polarity induce illusory edges that are mainly of the opposite polarity. Lines of either polarity (light or dark; Figs. VIII.107-1 and VIII.107-2) also induce illusory edges but, in this case, surprisingly, the illusory edge polarity depends on the direction of motion tracking, while with inducing edgelets (Figs. VIII.107-3 and VIII.107-4) it does not (Georgeson, 1980a). These judgments of polarity are sometimes difficult to make but are statistically reliable over a large number of trials.

Interactive Figure VIII.107-3 again allows you to test the graph-paper effect with different plaid angles, but here the plaids are formed from dark-to-light edgelets. The angular dependence appears to be similar to the case of thin lines (Interactive Fig. VIII.107-2). Also, in Interactive Figure VIII.107-3, as the angle increases from ±45° to ±85° from vertical, readers should also see the emergence of a striking illusion of *lightness*: each lozenge shape in the plaid appears much lighter than the one to its left, even though they are all identical. This lightness illusion does not occur at the smaller angles, even though all the plaids

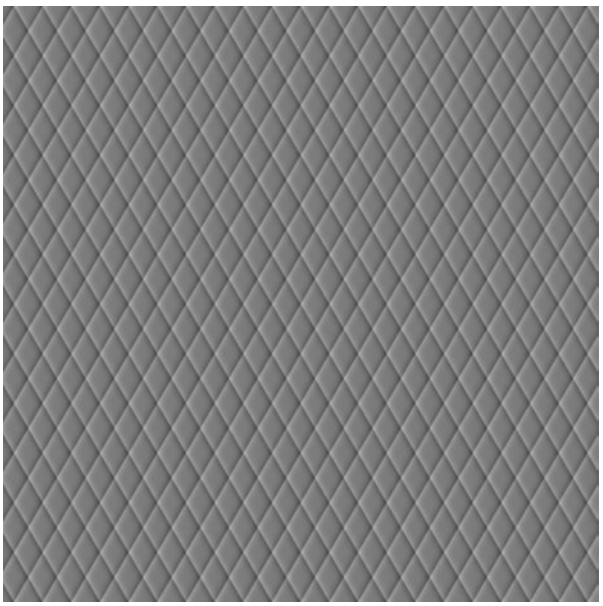


Figure VIII.107-3. Dark-to-light (DL) edgelets, medium contrast, ±30° from vertical. (Figure by Mark Georgeson, 2015.)

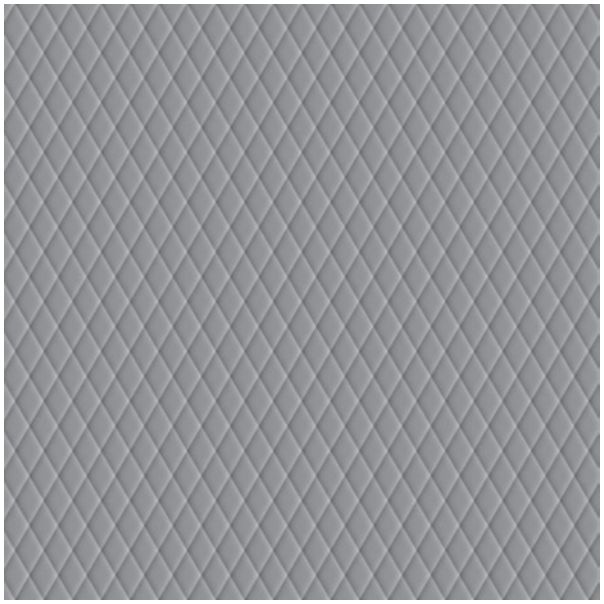
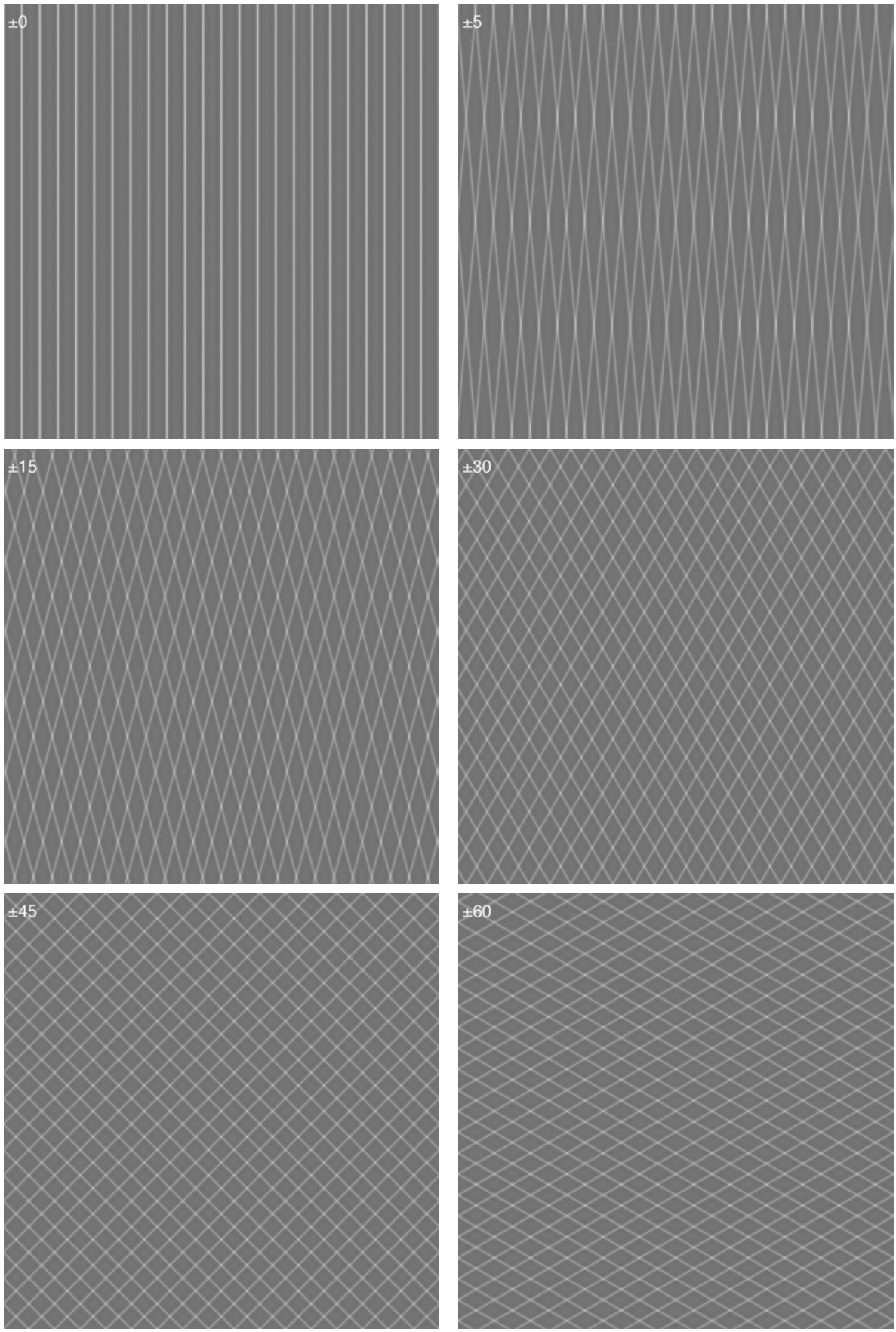
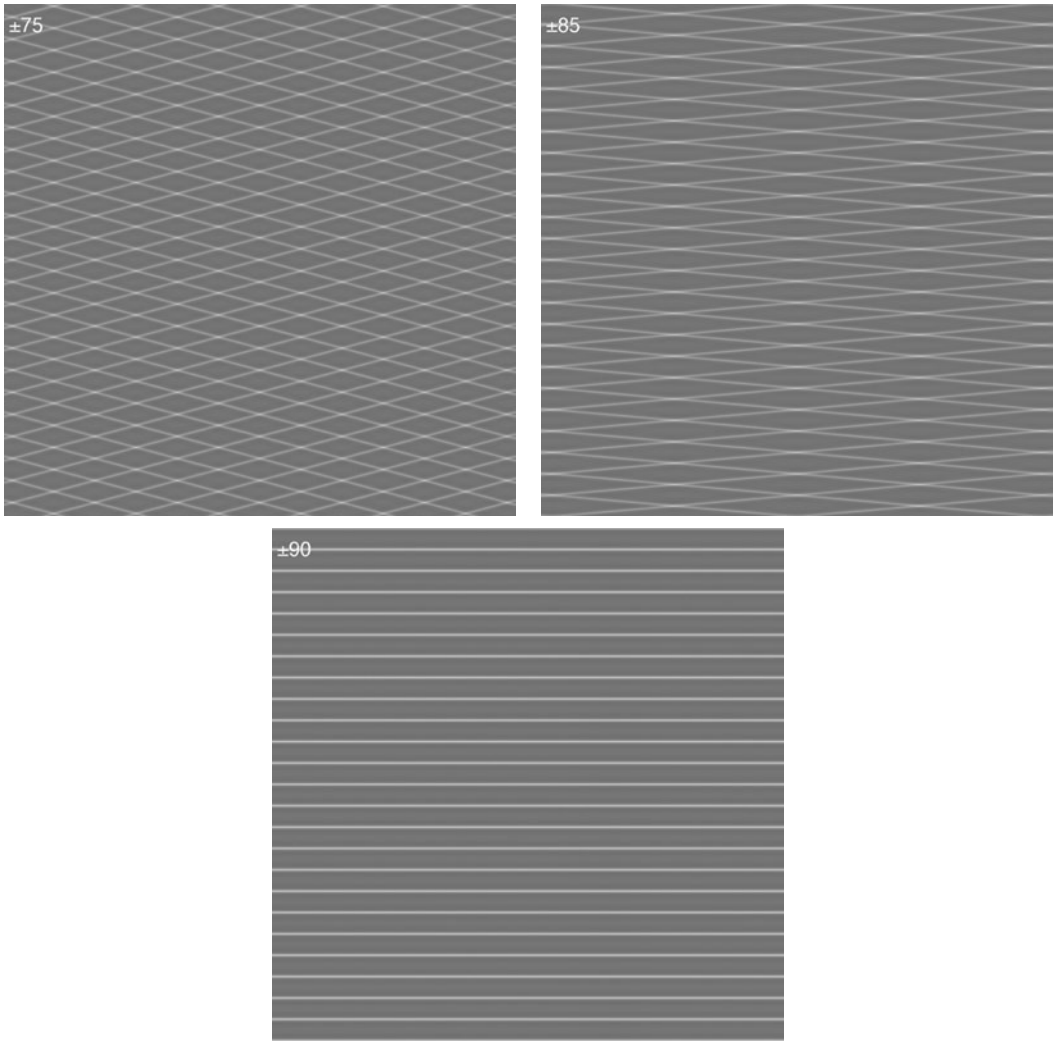


Figure VIII.107-4. Light-to-dark (LD) edgelets, medium contrast, ±30° from vertical. (Figure by Mark Georgeson, 2015.)



Interactive Figure VIII.107-2. Figure by Mark Georgeson, 2015.



Interactive Figure VIII.107-2. Continued

are the same, except for the angle between their component gratings. The illusion involves the "filling-in" or diffusion of brightness between edges and deserves further investigation but is beyond the scope of this chapter. A related phenomenon can be found on Michael Bach's excellent website of optical illusions (Bach, 2010).

Stereo Depth

A striking feature of the graph-paper effect is its appearance in 3D (Georgeson, 1980a). If you have normal stereo vision (about 5% of the population do not), then you should see that the illusory lines form a flat plane, parallel to the screen, floating in depth just about where your pen tip is. If you move the pen tip closer or further from the screen, this illusory plane should shift in depth accordingly. Advanced users may be able to converge their eyes *behind* the plane of the screen (a trick that is not easy to acquire), and then the illusory plane should appear behind the computer screen. Evidently when it is seen in depth, it is seen at the distance where the lines of sight of the two eyes converge. An object placed at this distance would project images to the same retinal positions in both eyes. In the language of stereo vision, these two images have "zero disparity." Some neurons in the visual cortex are specifically tuned to respond best to pairs of features with zero disparity, and so the illusion may be specifically linked to the activation of such cells. We return to this later.

Does the illusion appear as a plane because the inducing image is itself a plane in depth? If the inducing image had a 3D shape, would the illusion follow that shape? This is an interesting possibility, but the answer is "no." If we use real graph paper and fold it to create a "step" in depth, where the left half plane is nearer or further than the right half, then the illusion does not follow this shape. It remains as a plane that may appear nearer or further than the inducing planes, or even in between them, according to the convergence distance of the eyes (discussed earlier).

RELEVANCE TO MAINSTREAM VISION RESEARCH

Why should we care that people sometimes see things that are not there? Indeed, if such subjective effects were

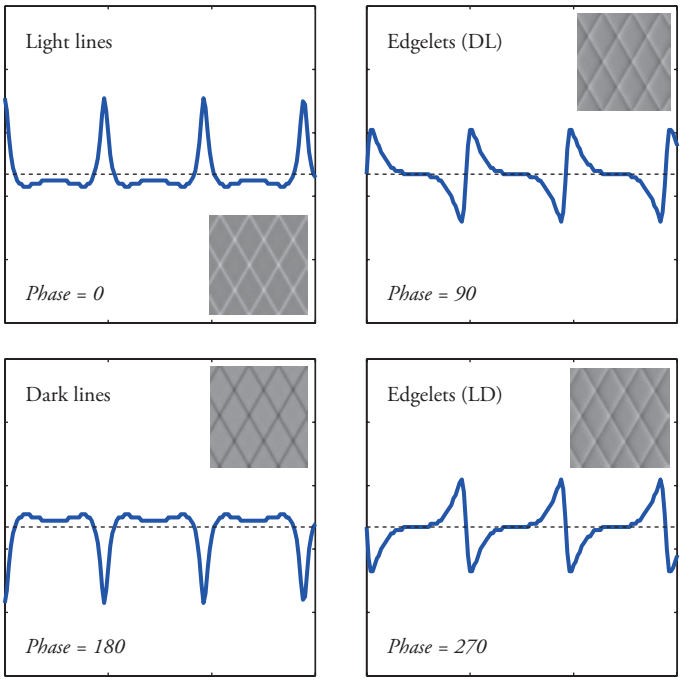


Figure VIII.107-5. Luminance profiles for the gratings that form the plaids in Figures VIII.107-1 to VIII.107-4 (and Interactive Fig. VIII.107-1). All four have the same Fourier amplitude spectrum, but their phases differ in steps of a quarter cycle (90° of phase). (Figure by Mark Georgeson, 2015.)

arbitrary and unsystematic, we probably should not care. They might arise from vivid imagination, biased reporting, compliance with the experimenter, or worse. But we should not ignore them merely because they are "subjective." Since *all* perception is ultimately a private experience that we cannot share with others, it would be illogical to dismiss some perceptual experiences as more subjective than others. In fact, there is a long and distinguished history of investigation throughout the 19th and 20th centuries into subjective phenomena, especially those evoked by spatially uniform flickering lights, in the expectation that they may shed new light on the mechanisms of vision. From the perspective of visual science, what really matters is whether their form and occurrence are systematically related to physical stimulus variables that can be measured and controlled, for then the *psychophysics* of subjective phenomena becomes possible. The graph-paper effect passes this test, as do various related effects and aftereffects, discussed next.

RELATED VISUAL EFFECTS AND AFTEREFFECTS

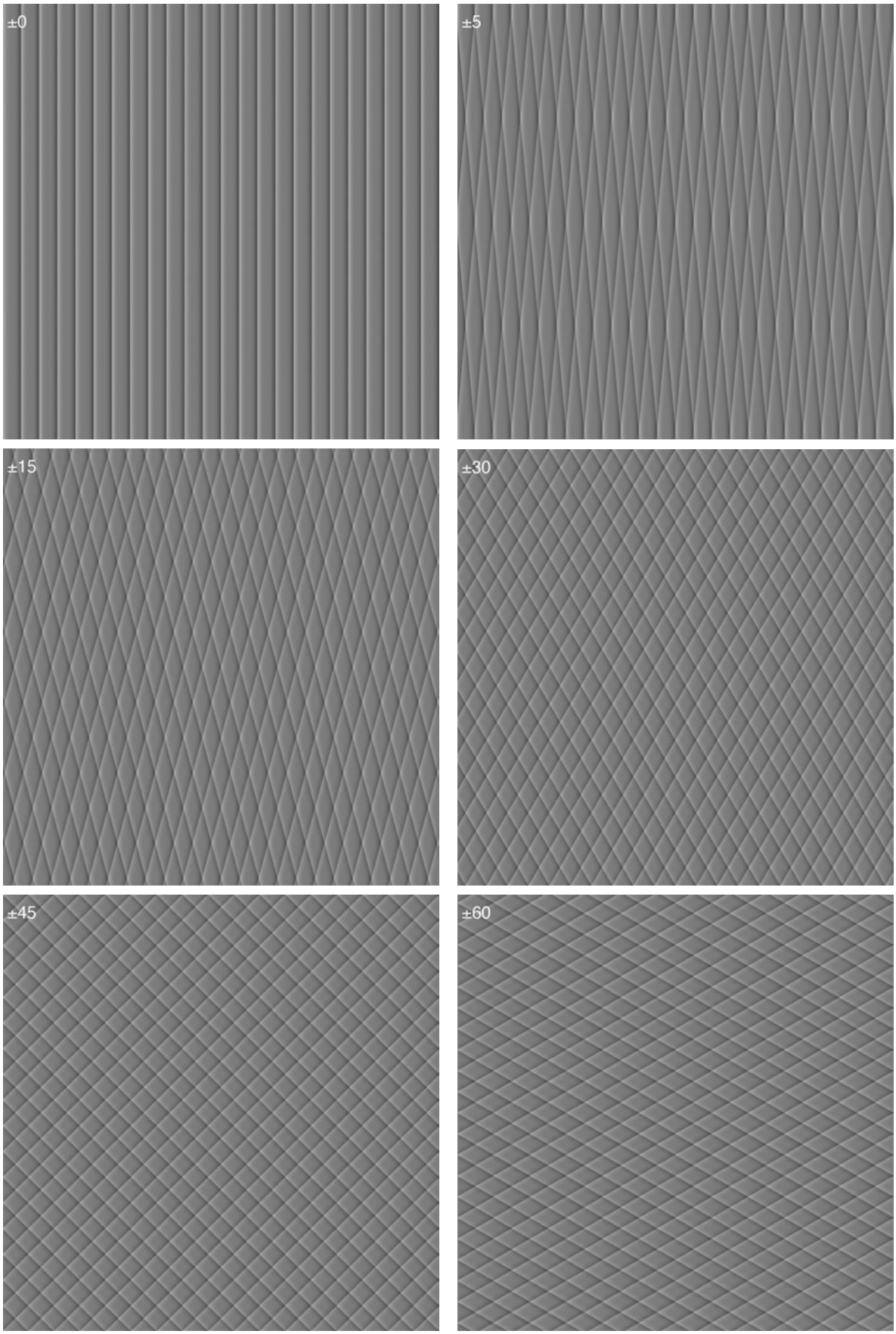
Flickering Lights

Smythies (1959) summarized the previous 150 years of work on the geometric patterns that can be evoked by large, spatially uniform fields flickering on and off at various rates from, for example, 2 Hz to 50 Hz. All the giants of 19th-century visual science (e.g., Purkinje, Fechner, Helmholtz, Brewster, Aubert, Exner) studied or wrote about these phenomena, and there is general consensus that large, fast-flickering fields of light can induce a wide range of fine, detailed, highly regular geometrical forms such as hexagonal figures, grids, radial lines, curved lines, mosaics, honeycombs, cobwebs, herring-bone, and so on. Systematic, complex movement—translation, rotation, expansion, oscillation—of these patterns is also widely reported. Smythies (1959) found that the patterns became finer and had smaller elements as flicker rate increased. Remole (1973, 1974) measured the threshold intensities of light required to evoke these patterns, as a function of flicker frequency in both monocular and binocular viewing. Thresholds were lowest at flicker rates of 10 Hz to 20 Hz. Diffuse patterns could be evoked at lower intensities, while the thresholds for geometric patterns were systematically higher and the thresholds for subjective movement or oscillation of these patterns were higher still. Interestingly, the thresholds for geometric patterns and movement were about twice as high with one eye as with two, implying that these phenomena arise at or after the stage at which signals from the left and right eyes are combined in the visual cortex. No such difference was found for the diffuse patterns, suggesting an earlier, monocularly driven site in the visual pathway.

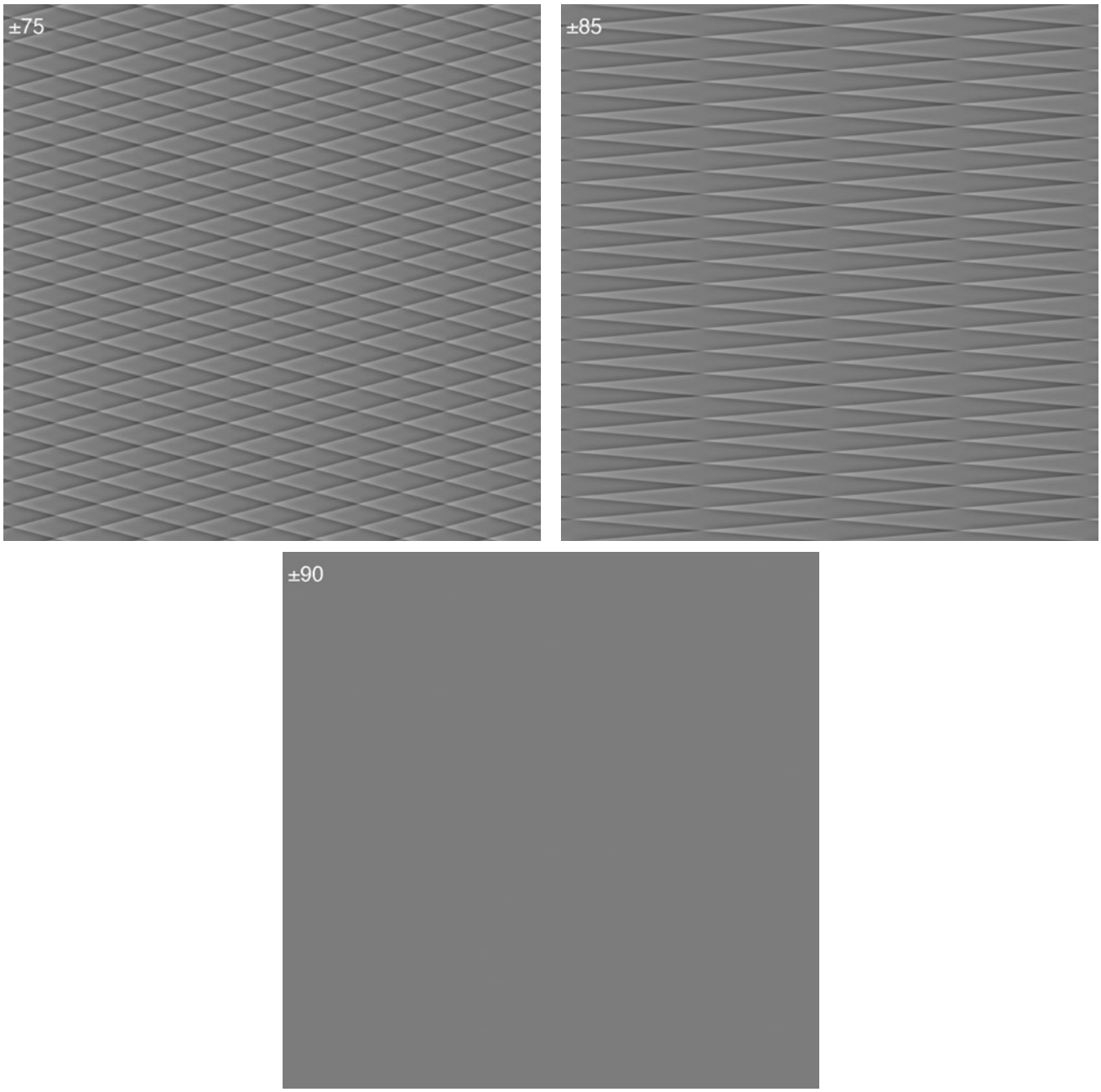
Aftereffects of Pattern Adaptation

In 1957 Donald MacKay famously described "MacKay's rays" in which an adaptation period (e.g., 30 s) spent viewing a pattern of high-contrast radial lines (the "rays") gave rise to a striking visual aftereffect of fast, rotary motion. On a completely blank white test screen, observers report seeing a chaotic field full of small, colorless particles or "grains of rice," streaming clockwise or anticlockwise or alternating in direction. Almost any high-contrast grating pattern—linear, radial, or concentric—can evoke the effect, and the subsequent illusory motion is always roughly at right angles to the pattern lines that induce it. Similar reports can be found in much earlier work (Hunter, 1915). Op artists in the 1960s incorporated these shimmering motions into their work (see Kumar & Glaser, 2006, for a recent experimental study). Like the flicker-induced effects, MacKay's rays suggest a functional link between distinct mechanisms for contour (spatial form) and movement. The idea of parallel systems for pattern and for motion was intensively studied in the 1970s and 1980s, drawing on the physiological and functional distinctions between X and Y cells in the cat's visual system and between P and M pathways in the primate visual system.

If adaptation to stationary contours somehow reveals or "unmasks" a response to motion, I reasoned that the reverse might also be true. And indeed, it turns out that adapting to patterns of rotating dots or blobs can evoke illusory radial lines or curves as an aftereffect (Georgeson, 1976a). If the blobs rotated in just one direction, the illusory lines appeared to rotate in the opposite direction (the classic motion aftereffect). When two blob patterns were superimposed, rotating in opposite directions, both motions were seen during adaptation, but they cancelled each other in the aftereffect; the illusory radial lines now appeared stationary (Georgeson, 1976a).



Interactive Figure VIII.107-3. Figure by Mark Georgeson, 2015.



Interactive Figure VIII.107-3. Continued

Finally, as well as evoking the appearance of streaming motion, adapting to stationary or flickering gratings at one orientation can reliably produce an aftereffect on a blank field in which "illusory gratings" are seen at other orientations or at different spatial frequencies. For example, adapting to a vertical grating produced (a) the illusion of a noisy grid or plaid of fine lines oriented about ±40° from vertical, and/or (b) the illusion of vertical gratings that were finer (higher spatial frequency) or coarser (lower spatial frequency) than the adapting grating (Georgeson, 1976b, 1980b). This suggests that when activity in orientation-tuned visual neurons is reduced by adaptation, the activity in other orientation-tuned neurons is increased—to such an extent that this spurious activity becomes visible, just as if those oriented contours were really present in the retinal image. But how might this indirect activation come about?

THEORY OF THE GRAPH-PAPER EFFECT AND RELATED ILLUSIONS

Inhibition and Disinhibition

Is there any general theory that might broadly account for the bizarre range of illusory effects we encountered here? I think there is, based on three main ideas: First, cells in the visual cortex of the brain are selectively tuned to

be responsive to different subsets of stimuli—to different orientations, sizes, binocular disparities, and directions of motion. This is what physicists might call the "standard model" of early visual processing, supported by the past 60 years of research in visual neurophysiology and psychophysics. These neurons are systematically arranged in columns or clusters of cells, such that adjacent columns are responsive to similar orientations or directions of motion (see Fig. VIII.107-6). Second, we suppose that there are mutual inhibitory or suppressive interactions between nearby neurons: cell A inhibits cell B, and cell B inhibits cell A. Inhibition between cells is ubiquitous in neural systems and serves many different functions in vision including spatial filtering, texture segmentation (Li, 2000), efficient coding, and noise reduction. Orientation-tuned inhibition at the cortical level has been proposed as the basis for geometric illusions of a more conventional kind, such as the Zöllner illusion and its variants (Blakemore, Carpenter, & Georgeson, 1970). Third, we suppose that cells can be activated by both direct and indirect routes—directly by excitation from other cells and indirectly by *disinhibition*: if B's activity *decreases*, then A's activity *increases* because it is inhibited less than before. Such patterns of inhibition and disinhibition have been observed at all levels of the nervous system, including the visual system, across many species. The compound eye of the horseshoe crab *Limulus* was famously studied as a model system for inhibitory interaction by Hartline and Ratliff (1957) who recorded and modeled mathematically both the inhibitory and disinhibitory effects. Let us consider three light-sensitive receptor units (A, B, C) of the *Limulus* eye. When unit C was activated by a spot of light, this inhibited a nearby unit (B), but the reduction of B's activity increased (disinhibited) the activity of unit A that was more remote from C.

Recent work in mammalian physiology has revealed analogous disinhibitory interactions in the visual cortex (e.g., Walker, Ohzawa, & Freeman, 2002), arising from the "suppressive surround" region that typically extends beyond the smaller "classical" receptive field of visual cortical cells. Corresponding psychophysical work has implied at least two forms of inhibitory effect at this level—called surround suppression and overlay suppression—and shown that the detectability of a visual target can be relieved from surround suppression via disinhibition (Petrov, Carandini, & McKee, 2005). In short, the three components (selective tuning, inhibition, and disinhibition) that we need to explain the illusory gratings are already available to us in the standard model of visual processing.

Explaining the Aftereffects and the Graph-Paper Effect

I suggest that the logic applied so elegantly by Hartline and Ratliff (1957) to the *Limulus* crab can also be applied to MacKay's rays, to the illusory grating aftereffects, and to the graph-paper effect. The inhibitory interactions take place in the visual cortex, where cells are selectively responsive to motion and pattern orientation, rather than in the retina, where they are not. Figure VIII.107-6 aims to clarify this idea by sketching a set of columns of orientation-tuned cells. In the upper row, the cells are selective for direction of motion at right-angles to their preferred orientation, as observed physiologically in the visual cortex of cats and primates. All adjacent cells are presumed to be mutually inhibitory (black connections).

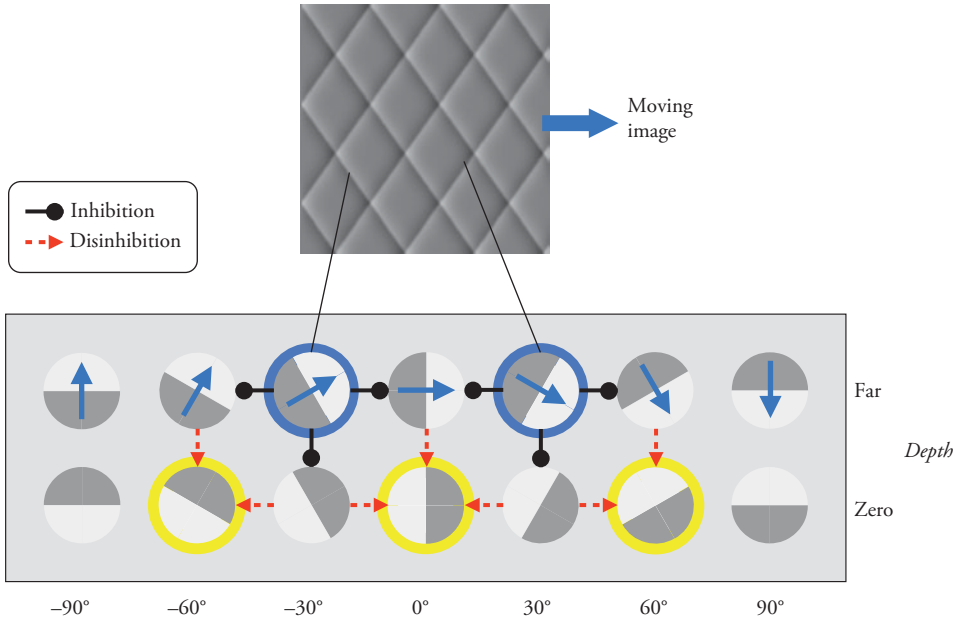


Figure VIII.107-6. Schematic neural model for the graph-paper effect, showing the inhibitory connections (black) and consequent disinhibitory effects (red), among a set of neurons or "detectors" tuned to different image properties—orientation, direction of motion, and stereo depth. Detectors ringed in blue are activated directly by the moving stimulus plaid; those ringed in yellow represent the illusory features, activated by disinhibition. Blue arrows (top row) denote selectivity for direction of motion; absence of blue arrows (bottom row) implies no motion selectivity. (Figure by Mark Georgeson, 2015.)

The streaming motion evoked by MacKay's rays arises when orientation-tuned cells that prefer slow or stationary contours are adapted (bottom row), thus releasing activity in motion-selective cells (top row). If this activity exceeds the background noise level by a reasonable margin, then motion will be rendered visible even though none exists on the blank screen. Its direction should always be at right angles to the adapting orientation(s), as observed.

The illusory grating aftereffects (Georgeson, 1976b, 1980b) occur when cells preferring a narrow range of orientations are adapted and cells at adjacent orientations are consequently released from inhibition. This activity would normally be caused by oblique lines or gratings, and so the perceptual system quite reasonably concludes that such contours are actually present. As with many other illusions, perceptual decoding (interpretation) is working as normal—business as usual—but the sensory evidence it has to work with has arrived by an unusual route.

These aftereffects arise from a one-step process of disinhibition (adapt B, release A). The graph-paper effect, on the other hand, is observed during stimulation, not as an aftereffect. It requires the two-step process described by Hartline and Ratliff (1957): stimulate C, inhibit B, release A. Figure VIII.107-6 traces these proposed steps. The oblique moving plaid activates motion cells in the upper row (ringed in blue); they inhibit all their neighbors. The consequent ripple of disinhibition (red arrows) carries through to cells tuned to vertical (0°) and to cells tuned to even more remote orientations (±60°; ringed in yellow). This explains the orientations seen in the graph-paper effect and might also explain why (for this plaid) vertical is the dominant illusory orientation: it is disinhibited via more routes than the remote (±60°) orientations, and so the activation may be stronger and more reliable. The illusory features appear to move with our gaze because (we presume) the cell activity that underlies them carries information about position and orientation but not about retinal motion. Hence these features are stationary in *retinal* coordinates, as discussed previously.

Taking this scheme (Fig. VIII.107-6) a step further, we can begin to understand the 3D nature of the effect. If the motion cells respond preferentially to far (or near) depths (top row), while the zero-disparity cells are less concerned with motion (bottom row), then presenting the plaid beyond the fixation distance, as we have done, will further enhance the selective activation of the motion cells and promote disinhibition of the zero-disparity cells. This explains why the illusion appears to lie in the plane of fixation—because that plane is specifically represented by the activity of cells that are sharply tuned for zero disparity.

What If We Do Not See These Effects?

Large numbers of people have seen the effects I have described and quantified them in experiments, but some do not see them or see them only weakly. I see them fairly strongly, perhaps through a combination of selective attention, extensive practice, and perceptual learning. It is even possible that, through perceptual learning, the inhibitory connections are strengthened and the illusory effects enhanced. On the other hand, it is not really surprising that such effects are generally rather faint. The sensory system would be dysfunctional if it frequently produced spurious signals that were mistaken for reality. Parameters and thresholds in sensory neural networks must surely be set to maximize the detection of reality and minimize illusion. Perhaps readers who cannot see these effects are more finely tuned to reality.

As Hartline and Ratliff (1957) remarked of the *Limulus* crab:

> The release of a receptor unit from the inhibiting effects of others, by causing those others to be inhibited by yet a third group of receptors, is interesting physiologically. Such "disinhibition" is not difficult to obtain, though it may require some pains to show a strong effect. (p. 371)

The same appears to hold in human vision for the graph-paper effect and its family of relations described here.

REFERENCES

Bach, M. (2010). Visual phenomena & optical illusions. http://www.michaelbach.de/ot/lum_diamond/index.html

Blakemore, C. B., Carpenter, R. H. S., & Georgeson, M. A. (1970). Lateral inhibition between orientation detectors in the human visual system. *Nature*, *228*, 37–39.

Georgeson, M. A. (1976a). Antagonism between channels for pattern and movement in human vision. *Nature*, *259*, 413–415.

Georgeson, M. A. (1976b). Psychophysical hallucinations of orientation and spatial frequency. *Perception*, *5*, 99–111.

Georgeson, M. A. (1980a). The graph-paper effect: Subjective stereoscopic patterns induced by moving gratings. *Perception*, *9*, 503–522.

Georgeson, M. A. (1980b). The perceived spatial frequency, contrast and orientation of illusory gratings. *Perception*, *9*, 695–712.

Hartline, H. K., & Ratliff, F. (1957). Inhibitory interaction of receptor units in the eye of limulus. *Journal of General Physiology*, *40*, 357–376.

Hunter, W. S. (1915). Retinal factors in visual after-movement. *Psychological Review*, *22*, 479–489.

Kumar, T., & Glaser, D. A. (2006). Illusory motion in Enigma: A psychophysical investigation. *Proceedings of the National Academy of Sciences USA*, *103*, 1947–1952.

Li, Z. (2000). Pre-attentive segmentation in the primary visual cortex. *Spatial Vision*, *13*, 25–50.

MacKay, D. M. (1957). Moving visual images produced by regular stationary contours. *Nature*, *180*, 849–850, 1145–1146.

Petrov, Y., Carandini, M., & McKee, S. (2005). Two distinct mechanisms of suppression in human vision. *Journal of Neuroscience*, *25*, 8704–8707.

Remole, A. (1973). Subjective patterns in a flickering field: Binocular vs monocular observation. *Journal of the Optical Society of America A*, *63*, 745–748.

Remole, A. (1974). Luminance thresholds for perceived movement in a flickering field. *Journal of the Optical Society of America A*, *64*, 1133–1135.

Smythies, J. R. (1959). The stroboscopic patterns: II. The phenomenology of the bright phase and after-images. *British Journal of Psychology*, *50*, 305–324.

Walker, G. A., Ohzawa, I., & Freeman, R. D. (2002). Disinhibition outside receptive fields in the visual cortex. *The Journal of Neuroscience*, *22*, 5659–5668.

PART IX
Adaptation

Chapter 108
Adaptation to Brightness Change, Contours, Jogging, and Apparent Motion

Stuart Anstis

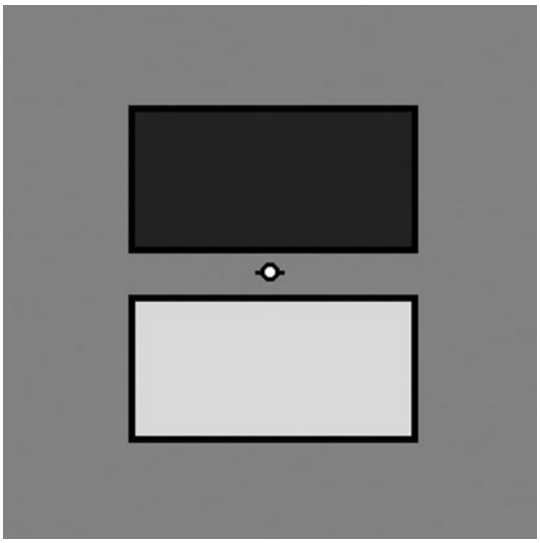
Frisby and Stone (2010) have dubbed adaptation the "psychophysicist's electrode," and John Mollon once (1974) famously said, "If it adapts, it's there." Psychologists piously hope that their many experiments on visual adaptation (Thompson & Burr, 2009) will tell physiologists where to look inside the brain. This chapter describes visual adaptation to temporal ramps, spatial edges, and apparent motion (AM) and touches on kinesthetic aftereffects from jogging.

SAWTOOTH ADAPTION

Sawtooth adaptation (or a ramp aftereffect) is produced by gazing at a spatially uniform patch whose luminance is temporally modulated by a repetitive sawtooth, either gradually dimming and turning sharply back on (rapid-on) or gradually brightening and turning sharply back off (rapid-off). This can lead to two different kinds of aftereffect: a shift aftereffect (ramp aftereffect), in which a steady test patch appears to be gradually brightening or dimming in the opposite direction to the adaptor, and a threshold-elevation aftereffect, in which a weak sawtooth of the same polarity as the adaptor becomes more difficult to see. I describe shift aftereffects (ramp aftereffects) first. It is well known that adapting to a static, dark (or light) square gives rise to a light (or dark) negative afterimage. I also discovered a dynamic version, namely adaptation to gradual change of luminance. This can be called a ramp aftereffect (Video IX.108-1).

Following adaptation to a spatially uniform patch of light that is growing gradually brighter, a subsequently viewed static gray test patch appears to be growing gradually dimmer (Video IX.108-1, upper panel). Conversely, following adaptation to a patch that is growing gradually dimmer, a gray test patch appears to be growing gradually brighter (lower panel; Anstis, 1967). A single rising or falling adapting ramp is usually sufficient to produce a detectable aftereffect, but it is convenient in practice to fold the ramp into a repetitive sawtooth. Typically, an adaptation period might consist of 30 sawteeth, each lasting 1 s during which the luminance would grow or shrink through 1 log unit. During the aftereffect, the test stimulus appears to brighten or darken monotonically for 3 to 8 s. In Figure IX.108-1 the solid black lines show the luminance profile of a typical stimulus over time in (a) and (c) and the perceived aftereffect in (b) and (d). Aftereffects are more conspicuous if either the adapting or the test field—or both—are spatially shaded from light to dark (not shown), which produces an impression of motion when the adapting field changes in luminance or when an aftereffect changes the subjective brightness of the test field.

Effects can be measured by a matching method. The maximum aftereffects, equivalent to about ± 15 dB/s, can be obtained from a large adapting amplitude of ± 20 dB/s (=1 log unit/s), but they are fairly insensitive to the temporal rate of this sweep. The channels are present out to an eccentricity of at least 40°, but they almost disappear during dark adaptation. The ramp aftereffects were asymmetrical: the subjectively darkening aftereffect produced by a brightening adapting ramp were 30% stronger than vice versa (Arnold & Anstis, 1993), suggesting that units that detect gradually rising luminance are more plentiful than units that detect gradually falling luminance. Aftereffects



Video IX.108-1. Ramp aftereffect. The upper adapting panel brightens repetitively while the lower adapting panel dims repetitively, both driven by sawtooth waveforms. When the panel luminances are briefly held constant, the upper panel appears subjectively to dim while the lower panel appears to brighten. (Figure by Stuart Anstis.)

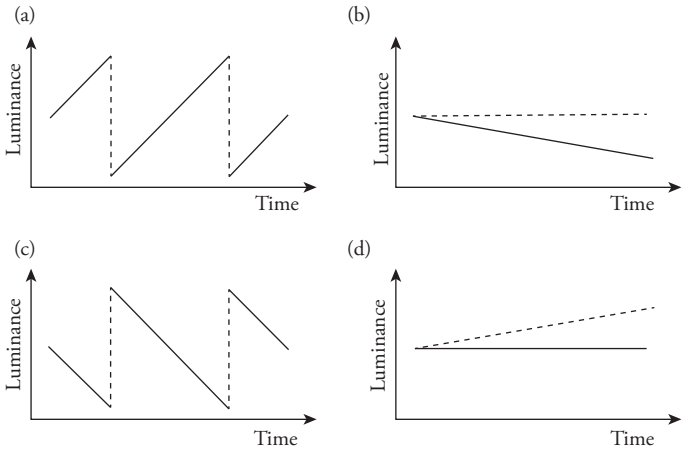


Figure IX.108-1. (a) Adaptation to a spatially uniform patch that repetitively brightness along a sawtooth luminance waveform shows an aftereffect, in which (b) a steady test patch of light appears to grow gradually dimmer. Conversely, adapting to a patch that (c) repetitively dims, yields an aftereffect in which (d) a test patch appears to grow gradually brighter. (Figure by Stuart Anstis.)

also decline markedly at low (scotopic) luminance levels, falling from 12 to 3 dB/s when the luminance is reduced by 2.5 log units. The spatial resolution of the underlying temporal channels that create the aftereffects was very low, estimated to be equal to an acuity of only 20/400. Thus the "perceptive fields" underlying the aftereffect are extremely coarse, and they rapidly become even coarser with increasing eccentricity (Anstis & Harris, 1987). Bosten and Macleod (2013) report a surprising relationship between static and dynamic luminance. After adaptation to a brightening (rapid-off) sawtooth, the threshold is raised for detecting a static disk that is slightly lighter than its surround (spatial increment).

Sawtooth adaptation can coexist with simultaneous contrast; if a small grey spot of constant luminance is centered in a brightening field, it appears to be growing gradually dimmer by simultaneous contrast. This apparent (illusory) dimming leads to a brightening aftereffect in the spot (Anstis, 1979). Chromatic sawtooth aftereffects have been sought but not found. If a dimming red patch is superimposed on a brightening green patch, the result is an equiluminous patch that modulates from red through yellow to green, but adaptation to this yields no aftereffect of apparent color change. This lack of color response, plus the fact that sawtooth aftereffects show no interocular transfer (Anstis, 1967), suggest that the underlying ON and OFF channels lie early in the M pathways, which respond to flicker but are believed to be insensitive to colors, unlike P pathways. Ramp aftereffects imply that neural units exist selective for gradual brightening and for gradual dimming. In a simple model, a summation unit receives two inputs from a single retinal region: a direct input and a delayed input that signal, respectively, the present and the immediately past luminance. Subtracting the delayed from the direct input gives zero signal for an unchanging luminance, a positive signal for a brightening field, or a negative signal for a dimming field. This model is equivalent to Reichardt's (1961) motion detector but with zero spatial shift. Reichardt's motion detector receives inputs from two adjacent retinal regions and detects motion from one region to the other, whereas the dimming/brightening model receives two inputs from the same retinal region and effectively yields the time derivative of the luminance.

There is a related effect of apparent brightening when gratings with an asymmetrical ramp (sawtooth) waveform move across the retina (Cavanagh & Anstis, 1986). If the grating moves in a direction that stimulates a given retinal area with a gradual brightening, interspersed with sudden darkening, the grating looks apparently lighter while it moves. When it moves in the opposite direction, it looks apparently darker (Fig. IX.108-2). Or we can move our finger across the figure and track that. Some people also report faint colors associated with the brightness changes.

The movement drifts the spatial sawtooth across retinal receptors, which therefore see a temporal sawtooth. The apparent brightness change bespeaks a nonlinearity in the visual system that under-reports the rapid phase and/or over-reports the slow phase of the sawtooth.

Adapting to a temporal sawtooth can also raise the threshold for test sawteeth of the same polarity. Exposure to a large uniform field modulated in luminance by a sawtooth function, repeating between one and five times per second, raised the threshold for detection of a test stimulus of similar waveform by a factor of 2 to 4.5. In comparison, the threshold elevation for a test stimulus of the inverse waveform was only half as great. This polarity-sensitive adaptation indicates that separate visual channels signal "brightening" and "darkening" (Hanly & Mackay, 1979; Purkiss & DeMarco, 2002). It is assumed that rapid-on sawtooth and rapid-off sawtooth tests are detected differentially by temporal ON and OFF channels, respectively. The visual system is generally more sensitive to rapid-off sawtooths than to rapid-on sawtooths, which strengthens the evidence that the visual system is more sensitive to sudden decrements than increments in light level (Bowen, Pokorny, & Smith 1989). Kremers, Lee, Pokorny, and Smith (1993) explored the physiology of temporal ON- and OFF-center channels in macaque retinal ganglion cells. They concluded that human luminance sawtooth sensitivities measured by psychophysics are consistent with selective detection through M pathways in the macaque visual system.

CONTRAST ADAPTATION IS PATTERN-SPECIFIC

It is well known that adapting to a white (or black) square gives rise to a negative afterimage that is black (or white). Adapting to a square that flickers between black and white on a grey surround gives no visible afterimage, provided that the time-integrated black and white are equal to the luminance of the grey surround. But it does leave a region of reduced sensitivity, so that a congruent low-contrast test square can become invisible. Webster and Mollon (1994) describe this as "contrast adaptation" and conclude that light adaptation adjusts sensitivity to mean luminance, while adaptation to contrast (i.e., to flicker) adjusts sensitivity to variations in luminance.

Note that contrast adaptation absolutely requires that the adapting and test regions must be congruent (Anstis, 2014). This is shown in Video IX.108-2, in which the test

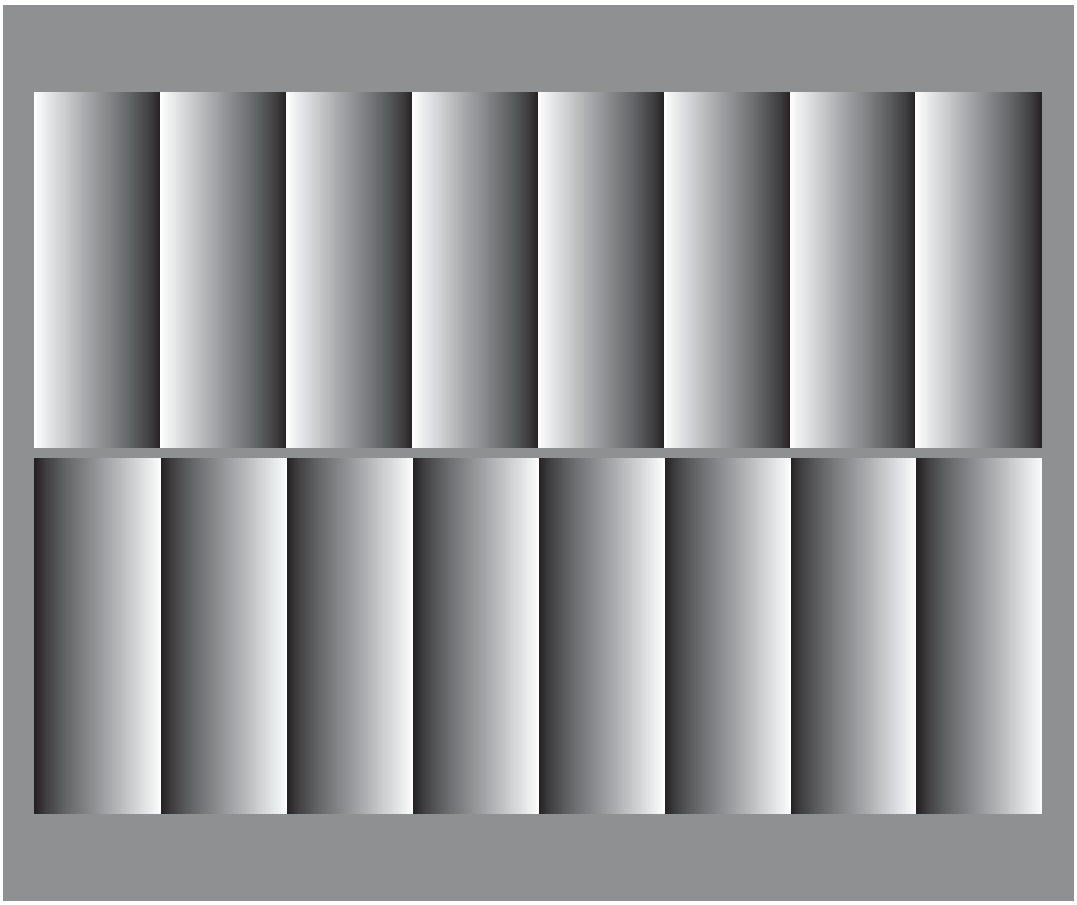
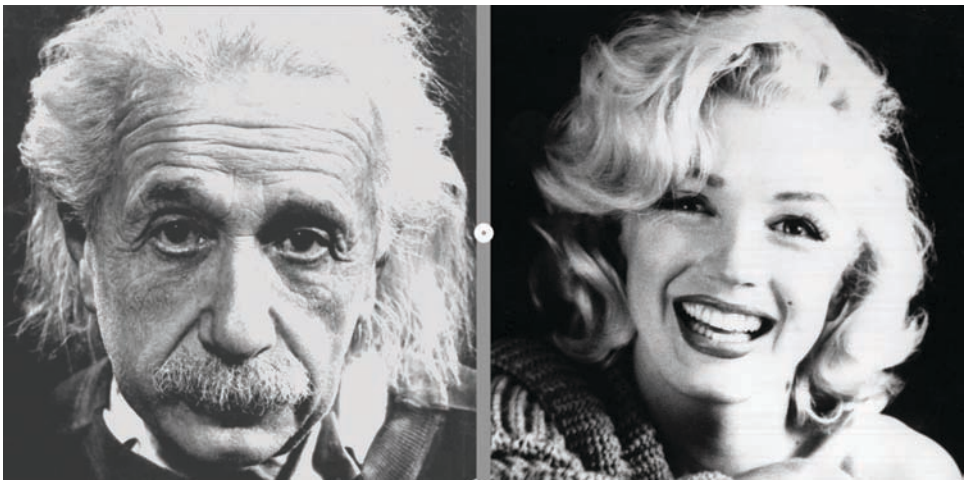


Figure IX.108-2. Gratings with sawtooth luminance profiles. If this picture is moved sharply to the left, the upper grating appears briefly dimmer and the lower grating appears briefly brighter. (Figure by Stuart Anstis.)

stimulus consists of two transparently superimposed, low-contrast grey-scale photos, one of Albert Einstein and the other of Marilyn Monroe (Oliva & Schyns, 1997). Two identical Einstein+Marilyn photos are set up side by side with a fixation point between them; each looks like a confused jumble, and neither face can be seen clearly. The adapting stimuli are high-contrast flickering versions of the two single components: Einstein on the left and Marilyn on the right. The resulting is adaptation makes Einstein fade out subjectively from the left-hand Einstein+Marilyn, which now looks like Marilyn. Conversely, Marilyn subjectively fades out from the Einstein+Marilyn on the right, which now looks like Einstein. This adaptation selectively picks out (and degrades) the test photo with which it is congruent



Video IX.108-2. Two identical test stimuli show Einstein on top of Marilyn Monroe in low contrast. Following adaptation to a flickering Einstein on the left and Marylin on the right, the (identical) test stimuli look like Marilyn on the left and Einstein on the right. (Figure by Stuart Anstis.)

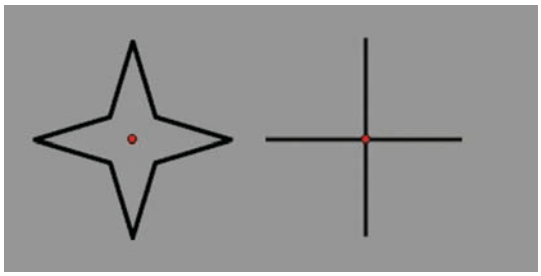
and has little effect on the other, superimposed but noncongruent test photo.

CONTOUR ADAPTATION

Contour adaptation is a process of reducing a shape's perceived contrast, or even making it invisible, by adapting to its flickering outline. Video IX.108-3 shows four low-contrast stars on a grey background. Two stars are slightly lighter than the background and two are darker. This is the test stimulus. While playing the video, fixate on the central spot. What is seen during the adaptation phase is outline contours of the top and bottom stars that flicker between black and white. This gives "contour adaptation" that reduces contrast sensitivity. The result is when one views the four test stars again, after the flickering adapters, one will probably be able to see only two. So adaptation has made two of the test stars invisible. Thus adapting to just the outline of a star makes the whole area of the test star invisible. This suggests that brightness information may be stored in edges and normally fills inward from them.

One can selectively adapt different contours within a stimulus. Video IX.108-4 shows two identical stars that are divided into light and dark gray quadrants. Adapt separately to each star by gazing at each fixation spot on different trials. Adapting to the cross deletes the test star's internal structure so that it appears uniform gray. Adapting to the external star contour deletes the test star's external structure so that it looks like four disembodied quadrants.

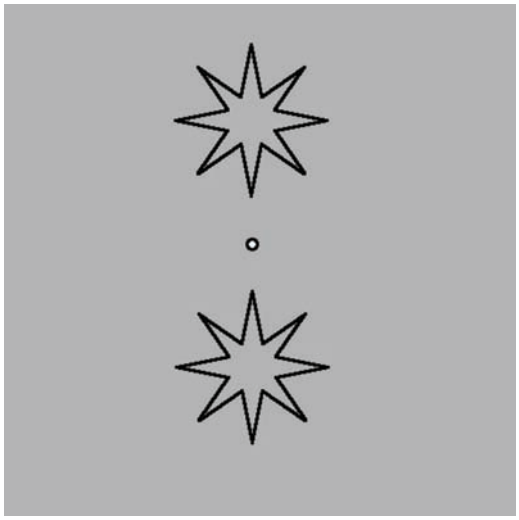
Contour adaptation suggests that a uniform light-colored shape is like a lake of brightness with a dam around its edge. Thus brightness information is stored in contours and normally fills-in from them. Adapting to a contour perceptually deletes it and breaches the dam, allowing the background brightness to fill-in like water, making the shape indistinguishable from the background. Filling always seems to work inward, from large surrounds into small enclosed objects, never outward (the same is true for simultaneous contrast). The reasons for this have not yet been studied.



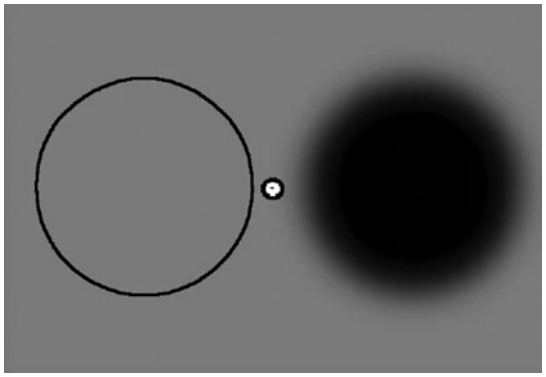
Video IX.108-4. Contour adaptation. The test stimuli are two four-pointed stars, each divided into light and dark grey quadrants. Adapt first to the thin outline that defines the left-hand star's outer edge. Result: The grey test star looks like four disembodied quadrants. Next, adapt to the cross that defines the inner structure of the right-hand star. Result: This test star looks like a uniform grey. (Figure by Stuart Anstis.)

Video IX.108-5 shows that edges are more important than surfaces. There are two round adapting shapes. On the left is a flickering outline circle, and on the right is a flickering blurred disk, which contains virtually all of the information of a disk except for the outline. In other words, the left and right adapting circles specialize in high versus low spatial frequencies. Adaptation to the flickering outline circle makes its whole low-contrast grey test disk disappear perceptually, while adaptation to the blurred flickering disk does not. Thus the outline circle is a much more efficient adaptor than the blurred disk. The fact that adapting out the edges makes the whole disk disappear confirms that brightness is coded at the edges and fills into the center from there.

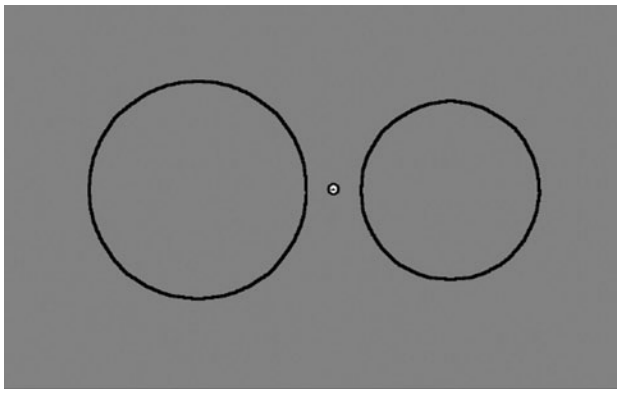
Video IX.108-6 shows two identical light-grey test annuli, alternating over time with two flickering adapting outline circles that are of slightly different sizes. The



Video IX.108-3. Contour adaptation. Adapting to the thin flickering outline contour of a star makes the entire low-contrast test star disappear. It makes no difference whether the test stars are lighter or darker than the surround. Not shown: A higher-contrast test star, say of Michelson contrast C, is perceptually reduced to a subjective contrast of C/2. (Figure by Stuart Anstis.)



Video IX.108-5. Contour adaptation. Adapting to the thin circular contours on the left is much more effective than adapting to the flickering blurred surface on the right. Following adaptation, the left test disk disappears while the right test disk remains visible. (Figure by Stuart Anstis.)



Video IX.108-6. Contour adaptation. Test stimuli are two identical light grey annuli. The small adapting circle on the right erases the inner edge of the test annulus, which fills in to look like a uniform light grey disk. The large adapting circle on the left erases the outer edge of the test annulus, which merges with the surround so that the round area enclosed by the annulus looks like a small, dark grey disk. (Figure by Stuart Anstis.)

adapting circle on the left is congruent with the outer edge of the test annulus and the circle on the right with the inner edge. Following adaptation, the two test annuli look very different. The annuli on the left and right look like a small dark disk and a larger light disk, respectively. This is because the adapting circles selectively erase the inner and outer edges of the test annulus, allowing the neighboring greys to fill-in. Contour adaptation is a useful technique for temporarily erasing contours to explore the effects of filling-in.

AFTEREFFECTS FROM JOGGING

Running is not a clockwork reflex but an adaptive skill. Runners constantly adjust their running to the local terrain, shortening their stride uphill, picking their way over rough ground, and so on. These adaptive processes are multimodal and can manifest themselves as both kinesthetic and visual negative aftereffects (see reviews by Herman, Grillner, Stein, and Stuart, 1976, and Patla, 1991). A useful test activity to demonstrate kinesthetic aftereffects is jogging in place (running on the spot). Adaptive changes in gait, namely negative aftereffects, occur after running on an exercise treadmill, a conveyor belt device that carries the runner backward at the same speed as he or she runs forward (Anstis, 1995). A runner adapted to the treadmill by running on it for 60 s at a speed of 8 km/h, with eyes closed (to rule out visual components such as optic flow) and holding on to the handrail (to avoid being flung off backward). When he dismounted from the treadmill and walked or ran forward on the ground, he had a subjective illusion of being apparently borne forward, as if on invisible wheels. More objectively, when asked to dismount from the treadmill and jog in place with eyes closed, a runner actually jogged forward through about 150 cm in 15 s without being aware of it. Running backward or sideways produced an aftereffect of an inadvertent backward or sideways drift. Running on a turntable that rotated clockwise about a vertical axis that ran through the middle of the runner's body required the runner to keep turning counterclockwise in order to keep facing north, and this led afterward to an inadvertent counterclockwise drift. The aftereffects dissipated gradually over time; if the runner stepped off the treadmill after 60 s and then jogged in place, either immediately or else after standing still for 15, 30, 60, or 120 s, with eyes closed throughout, the inadvertent forward motion was greatest for zero delay and decreased as the delay was progressively increased to 120 s. Thus the stored aftereffect gradually dissipated over the course of a minute or two.

A hopping technique helps to show the neural site of these aftereffects. The runner hopped on his preferred leg on the treadmill for 30 s at a speed of 5 km/hr. This slow speed was used because hopping is very fatiguing. The runner then dismounted and hopped in place on solid ground with eyes closed, using either the same (adapted) leg or the other (unadapted) leg on different trials. An aftereffect was found when the same leg was used (showing bipedal gait was not necessary) but *not* when the other leg was used. This failure of the aftereffect to transfer across legs suggests a peripheral rather than a central neural site. Thus the adaptation could not be visual or vestibular but must occur in those parts of the nervous system that control each leg separately. This aftereffect from jogging was purely kinesthetic with no visual component, since it was found when the observer's eyes were kept closed throughout, and it was also found for three congenitally blind observers, who bravely ran on the treadmill and upon dismounting showed the same inadvertent forward drift as normal observers. Durgin and Pelah (1999) confirmed that the running-in-place aftereffect may result from a recalibration of visuomotor control systems that takes place even in the absence of visual input. However, running is a multimodal activity, and visual components may also adapt during running (Durgin et al., 2005). Simon Levay, a neurophysiologist and long-distance cyclist, has observed a pedaling-contingent visual aftereffect (personal communication). When he mounted a *stationary* exercise bike and started pedaling, he noticed that the visual world momentarily seemed to move forward, in the direction of his intended movement.

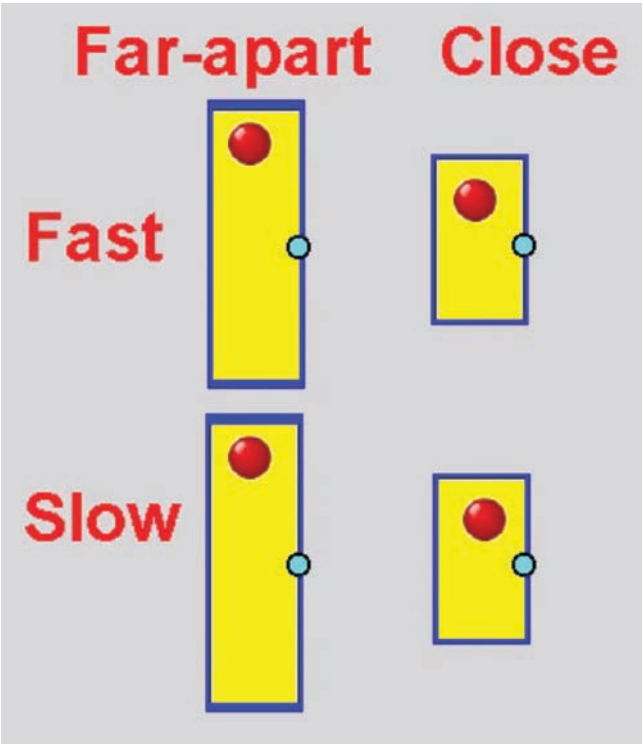
Visual feedback can also interfere with kinesthetic aftereffects: Durgin and Pelah (1999) reported that aftereffects were not limited to treadmill running but could also be strongly generated by running behind a golf cart when the eyes were closed but not when the eyes were open. Walking or running uphill can measurably affect judgments of slope. Walking up a slope of 8° gave an aftereffect that made a horizontal test path feel downhill and made an uphill test slope of 5° feel horizontal. What is the neural site of the kinesthetic aftereffects? Since the two legs adapted independently in hopping trials, we can rule out central components such as vestibular adaptation—the aftereffects occurred not necessarily within the leg itself but at least within the neural pathways that control each leg separately. Since running around the block was found to produce no significant aftereffects, the adaptation must specifically counteract the backward motion of the treadmill. We conclude that running on the treadmill adapts those neural pathways that compare the muscular

effort of swinging each leg forward (or backward) with the resulting postural position of each leg (Houk, Keifer, & Barto, 1993). The backward motion of the treadmill produces an artificial mismatch between motor output and normal postural feedback, for which the adaptation compensates or nulls out by adjusting internal gain parameters to bring output and feedback back into line. But once the runner steps on to solid ground, these newly adjusted parameters become inappropriate and manifest themselves as an aftereffect, which dissipates as the parameters automatically update to match the solid ground. Visual feedback also plays an important if not yet fully understood role. So these aftereffects, both kinesthetic and visual, reveal the continuous neural recalibration of the gait control system.

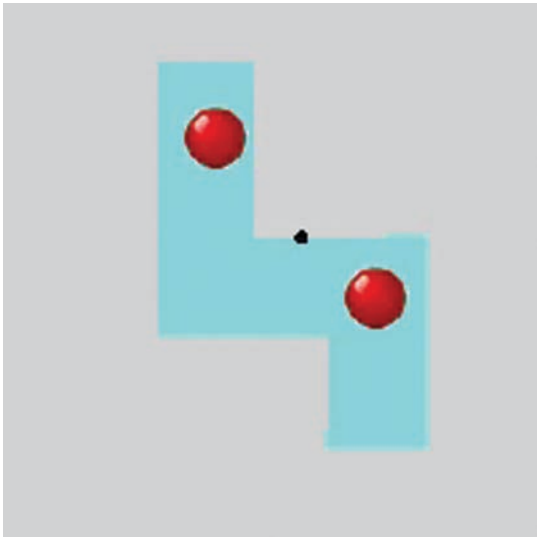
ADAPTATION TO APPARENT MOTION

A spot alternating between two positions can produce AM or stroboscopic motion. Video IX.108-7 shows that following prolonged inspection, the AM degenerates into flicker. Blue fixation spots are provided.

This adaptation effect depends on spacing and timing; the probability of seeing motion declines linearly with log spatial separation (over a range from 0.1° to 1°) and with log alternation rate (over a range from 2 to 4.5 Hz). So motion adapts out sooner in Video IX.108-7 for spots that are further apart (left-hand column) and alternate faster (top row). Adaptation to real motion strongly suppresses AM (not shown), which suggests that AM must be stimulating the same neural pathways as real motion.



Video IX.108-7. Adaptation to apparent motion. During strict fixation, prolonged apparent motion of a spot that jumps back and forth between two positions degenerates into a percept of two spots flickering in place. Somehow, adaptation causes loss of temporal phase information. Jumps with rapid alternations over large distances provide weak motion signals show the quickest adaptation and degenerate soonest. (Figure by Stuart Anstis.)

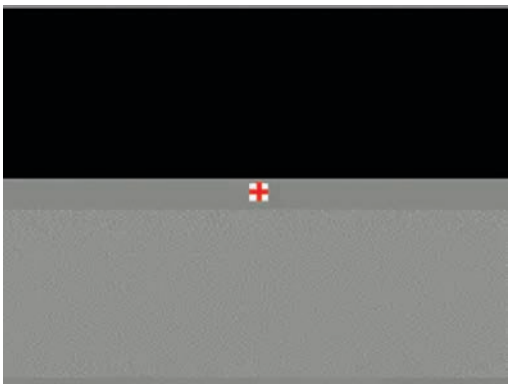


Video IX.108-8. Adaptation to apparent motion. The test spot jumps horizontally between two positions to the left and right of the fixation dot. The two adapting spots jump vertically. Although they land on the same two positions as the test spot, the vertical motions fail to adapt out the horizontal test motion. Thus the adapted motion signals are somewhat divorced from position. (Figure by Stuart Anstis.)

Video IX.108-8 shows that embedding the adapting spots in configurations of two vertically moving spot paths, a maneuver that alters the pattern of perceived adapting motion without altering the local retinal stimulation, minimizes the adaptation and has little effect on the horizontal path of the test motion. So the adapting mechanism must be responding to the path of seen motion and may be occurring in cortical area MT. This adaptation can be used to measure the strength of AM and shows that AM is strongest for small separations, low alternation rates, and high luminance contrast (Anstis, Giaschi, & Cogan, 1985).

ADAPTING TO FLICKER CHANGES APPARENT SPATIAL FREQUENCY

The contrast required to detect a flickering grating results from an interaction between its spatial and temporal frequencies. Robson (1966) showed that at low temporal frequencies, spatial contrast sensitivity is band pass, while at high temporal frequencies it is low pass. There is some evidence, though of a conflicting kind, that could attribute this performance to magnocellular pathways that are selectively sensitive to high temporal and low spatial frequencies (Goodbourn et al., 2012). Kaneko, Giaschi & Anstis (2015) reasoned that if such visual channels exist, then adapting to a patch of spatially uniform flicker would reduce the sensitivity of these channels, which would in turn affect their ability to code low spatial frequencies and



Video IX.108-9. Flicker adaptation. Following adaptation to the upper flickering panel, the upper test grating looks slightly finer than the lower one. Conclusion: Spatial frequencies are coded by a balance between transient/magno channels that respond to low-spatial, high-temporal frequencies and sustained/parvo channels that respond to high-spatial, low-temporal frequencies. The flicker is thought to adapt out the transient channels, reducing their low spatial frequency signal. (Figure by Stuart Anstis.)

therefore shift upward the apparent spatial frequency of a static coarse grating. And this is what we found. See Video IX.108-9, in which the upper half flickers at ~8 Hz. The two static test gratings are identical, but the upper one will look finer following adaptation.

Carrasco (1990) did an opposite experiment. She adapted her observers to a sinusoidal grating and tested them on a blurred flickering patch. She found that adapting to a low-spatial frequency grating *reduces* sensitivity to low temporal frequencies, while adapting to a high spatial frequency grating *increases* sensitivity to high temporal frequencies.

REFERENCES

Anstis, S. (2014). Pattern specificity of contrast adaptation. *i-Perception*, *5*, 73–74.

Anstis, S. M. (1967). Visual adaptation to gradual change of intensity. *Science*, *155*(3763), 710–712.

Anstis, S. M. (1979). Interactions between simultaneous contrast and adaptation to gradual change of luminance. *Perception*, *8*(5), 487–495.

Anstis, S. M. (1995). Aftereffects from jogging. *Experimental Brain Research*, *103*(3), 476–478.

Anstis, S., Giaschi, D., & Cogan, A. I., (1985). Adaptation to apparent motion. *Vision Research*, *25*(8), 1051–1062.

Anstis, S., & Harris J. (1987). Magnification factor for adaptation of a visual transient mechanism. *Journal of the Optical Society of America A*, *4*(8), 1688–1698.

Arnold, K., & Anstis, S. (1993). Properties of the visual channels that underlie adaptation to gradual change of luminance. *Vision Research*, *33*(1), 47–54.

Bosten, J. M., & Macleod, D. I. (2013). Mechanisms of the dimming and brightening aftereffects. *Journal of Vision*, *13*(6), pii: 11.

Bowen, R. W., Pokorny, & J., Smith, V. C. (1989). Sawtooth contrast sensitivity: Decrements have the edge. *Vision Research*, *29*(11), 1501–1509.

Carrasco, M. (1990). Visual space-time interactions: Effects of adapting to spatial frequencies on temporal sensitivity. *Perception & Psychophysics*, *48*(5), 488–496.

Cavanagh, P., & Anstis S. M. (1986). Brightness shift in drifting ramp gratings isolates a transient mechanism. *Vision Research*, *26*(6), 899–908.

Durgin, F. H., & Pelah, A. (1999). Visuomotor adaptation without vision? *Experimental Brain Research*, *127*(1), 12–18.

Durgin, F. H., Pelah, A., Fox, L. F., Lewis, J., Kane, R., & Walley, K. A. (2005). Self-motion perception during locomotor recalibration: more than meets the eye. *Journal of Experimental Psychology*: *Human Perception and Performance*, *31*(3), 398–419.

Frisby, J. P., & Stone, J. V. (2010). *Seeing: The computational approach to biological vision*. Cambridge MA: MIT Press.

Goodbourn, P. T., Bosten, J. M., Hogg, R. E., Bargary, G., Lawrance-Owen, A. J., & Mollon, J. D. (2012). Do different "magnocellular tasks" probe the same neural substrate? *Proceedings of the Royal Society B*, *279*(1745), 4263–4271.

Hanly, M., & MacKay, D. M. (1979). Polarity-sensitive perceptual adaptation to temporal sawtooth modulation of luminance. *Experimental Brain Research*, *35*(1), 37–46.

Herman, R. M., Grillner, S., Stein, P. S. G., & Stuart, D. G. (1976). *Neural control of locomotion*. New York: Plenum Press.

Houk, J. C., Keifer, J., & Barto, A. G. (1993). Distributed motor commands in the limb premotor network. *Trends in Neuroscience*, *16*, 27–33.

Kaneko, S., Giaschi, D. & Anstis, S. (2015). Flicker adaptation or superimposition raises the apparent spatial frequency of coarse test gratings. *Vision Research*, *108*, 85–92.

Kremers, J., Lee, B. B., Pokorny, J., & Smith, V. C. (1993). Responses of macaque ganglion cells and human observers to compound periodic waveforms. *Vision Research*, *33*(14), 1997–2011.

Mollon, J. D. (1974). After-effects and the brain. *New Scientist*, *61*, 479–482.

Patla, A. E. (1991). *Adaptability of human gait: Implications for the control of locomotion*. Amsterdam: Elsevier.

Reichardt, W. (1961). Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In W. A. Rosenblith (Ed.), *Sensory communication* (pp. 303–317). Cambridge, MA: MIT Press.

Robson, J. G. (1996). Spatial and temporal contrast-sensitivity functions of the visual system. *Journal of the Optical Society of America A*, *56*, 1141–1142.

Thompson, P., & Burr, D. (2009). Visual aftereffects. *Current Biology*, *19*(1), R11–R14. doi:10.1016/j.cub.2008.10.014.

Webster, M. A., & Mollon, J. D. (1994). The influence of contrast adaptation on color appearance. *Vision Research*, *34*(15), 1993–2020.

Chapter 109
The Color Dove Illusion

Chromatic Filling-In Effect Following a Spatial-Temporal Edge

Yuval Barkan and Hedva Spitzer

THE EFFECT

The classical afterimage (negative afterimage) occurs after the disappearance of a colored patch following a prolonged viewing. The effect induces an afterimage of a complementary color (e.g., yellow induces a bluish afterimage). In the color dove illusion, we observe a positive afterimage effect, as demonstrated in Figure IX.109-1 (Barkan & Spitzer, 2009, for an interactive version of the illusion). As the white dove flies across the sky with clouds, the color of the sky disappears (changes from red to white). At the same time, the uncolored bird captures the "reddish" color of its background. The dove's reddish apparent color persists for few seconds after the sky flashes to white. This effect shows that an afterimage color can linger in our visual perception and spread into an empty shape. The effect, of course, can be obtained from any empty shape. The color spread into the empty shape occurs only when the contour of the shape is maintained after the color disappears. This effect resembles a phenomenon that was previously investigated by Anstis, Rogers, and Henry (1978). The most important new aspect of the effect here is that it can be also obtained from a background that is spatially narrow and presented for a very short duration, (see the "flicker" option in Barkan & Spitzer, 2009). In this specific case, a narrow background flashes for a short duration (about 2 Hz) from white to red or to cyan. Accordingly, the dove obtains a pale color of the disappeared thin background.

VARIANTS AND PARAMETERS

The color dove illusion can be obtained with different colors and spatial surrounding regions, as can be seen in Figure IX.109-1 and Barken and Spitzer (2009). The effect occurs with any chromatic background. The perceived filling-in color inside the empty shape is similar to the inducing background color, but with a paler appearance (i.e. with less chromatic saturation).

The effect is also observed while there is only a thin chromatic ring around the empty shape (Fig. IX.109-1b). It can be seen that reducing the size of the chromatic ring causes a reduction in the effect (up to 20% reduction, according to preliminary experiments). In other words, drastically reducing the size of the inducing background does not cancel the effect. If the effect is derived from a combination of the simultaneous contrast and the afterimage effects (Anstis et al., 1978), we would expect that the color of dove illusion would be cancelled with such a narrow background (see later discussion).

The flickering variation of the color dove illusion (Barkan & Spitzer, 2009) shows that the effect can be perceived for a very short duration (in relation to the classical effect), that is, approximately 2 Hz. Preliminary experiments have shown that the effect does occur even with a very short duration (~300 ms) of the flickering background.

It appears that the effect can be manifested also with illusory contours (Kanizsa illusion) instead of real contours (for the empty shape), as can be seen in Figure IX.109-2. The possibility that illusory contours can substitute for real contours in regard to afterimages was previously suggested also for the filling-in afterimage effect (van Lier, Vergeer, & Anstis, 2009). It appears that the contour curvature and its length has an important role for the illusion prominence; however, this still needs to be investigated experimentally.

DISCUSSION

The color dove illusion is a positive aftereffect, which is different from classical positive aftereffects (Anstis et al., 1978). The special property of the effect is that it does not necessarily require a prolonged inducing duration of a thick induction background.

Anstis and colleagues (1978) investigated a resembling positive aftereffect. They suggested that the positive chromatic afterimage effect is induced due to the synergy of two known visual mechanisms (Fig. IX.109-3): simultaneous contrast and colored afterimage. These effects are briefly explained next.

The simultaneous contrast effect shows that the appearance of color or luminance of a patch depends not only on its luminance and color properties but also on the properties of

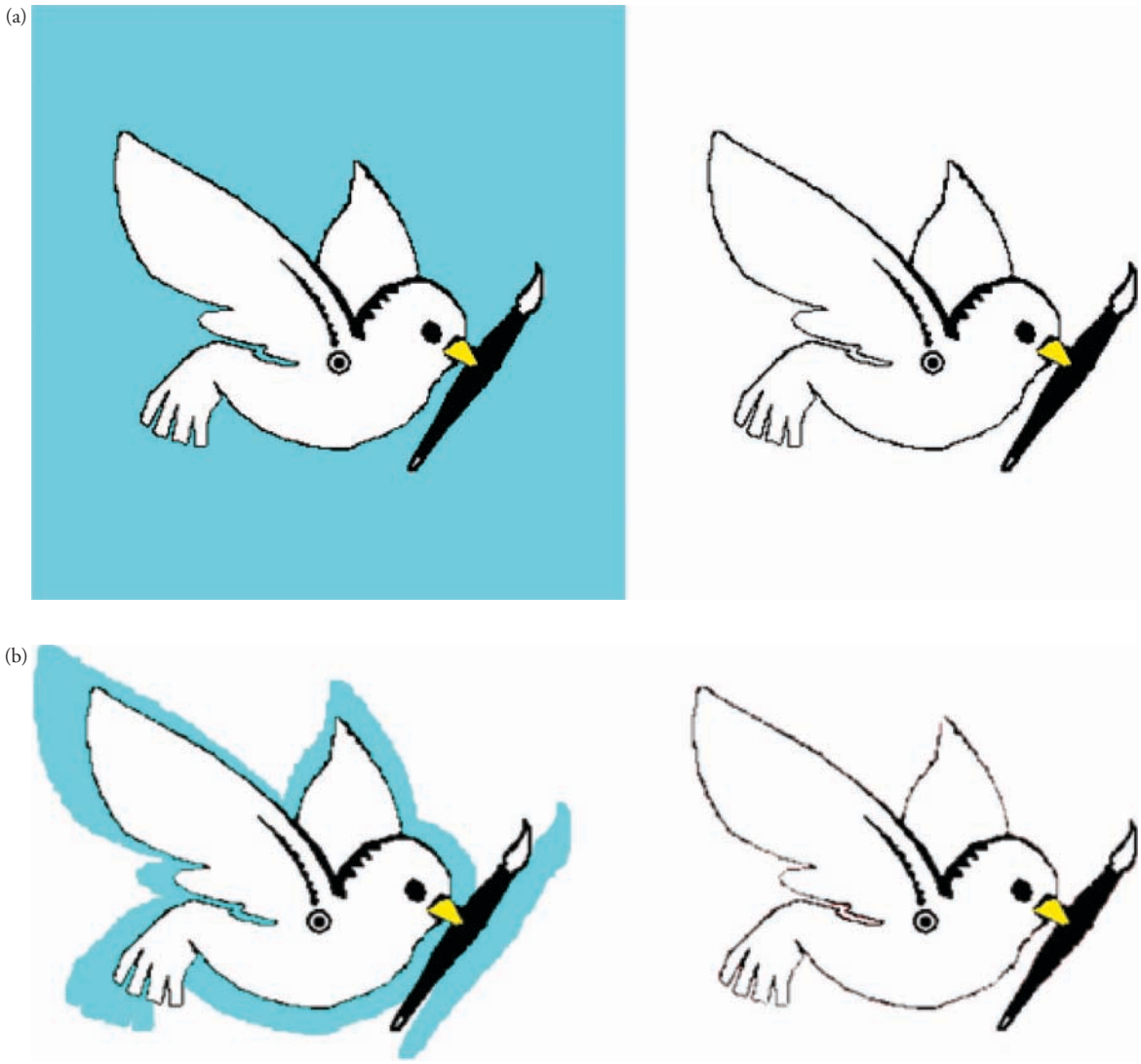


Figure IX.109-1. (a) An example of the color dove illusion. In order to see the effect, observers must fix their gaze on the central black point in the bird with the colored background (left side). After a few seconds, observers must move their gaze to the central black point in the bird without the colored background (right side). Observers should perceive that the "empty bird" is filled-in with a color similar to the background's chromatic color. The effect is perceived more prominently when the observer's gaze is steady and the background flashes, as can be seen in an interactive demonstration at http://illusionoftheyear.com/?p=1060) (b) A variant of the color dove illusion demonstrates that the effect does not require a large spatial area Similarly we have shown that the effect can be obtained with short induction time (see interactive demonstration referenced above).

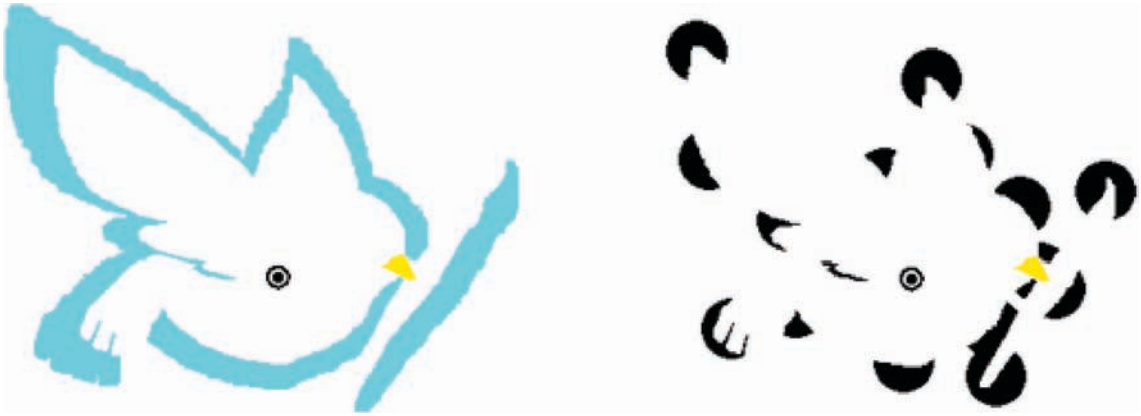


Figure IX.109-2. This example shows that the effect can also be obtained when the left contours do not consist of real edges but rather illusory contours, such as Kanizsa contours.

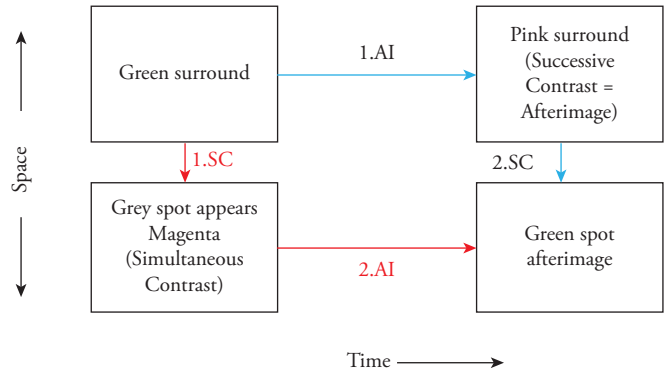


Figure IX.109-3. The two possible mechanisms described by Anstis et al. (1978), through the paths of two known visual effects. The explanation refers to a green inducting background, as originally described in Anstis et al. The red route: Simultaneous contrast (SC) first creates a magenta spot and then an afterimage (AI) causes the spot to appear green. The blue route: An AI creates the illusion of a magenta background that only then causes the neutral spot to appear green by SC.

its surrounding stimuli. An achromatic patch surrounded by a chromatic background will obtain a complementary color simultaneous to the presentation of the chromatic background.

"Afterimage effect" refers to an image continuing to appear in one's visual perception after the exposure to the original image has ceased. These two effects induce the perception of complementary color, in relation to a surround region chromaticity (simultaneous contrast) or in relation to the previous chromatic patch (afterimage). The classical simultaneous contrast effect requires a large inducing area (Blackwell & Buchsbaum, 1988; Semo, 2001; Sezikeye & Gurnsey, 2009; Tsofe, 2004; Zaidi, Yoshimi, Flanigan, & Canova, 1992), while the classical afterimage effect requires prolonged inducing duration.

In the current illusion we demonstrate that the effect is not abolished when stimulus requirements for both simultaneous contrast and classical afterimage are not fulfilled. This is demonstrated in Figure IX.109-1b (the flicker option), when the surround-colored background is thin (and therefore "blocks" simultaneous contrast) and is presented for a short duration (and therefore "blocks" the afterimage). The lack of abolishment of the effect in these stimuli parameters suggests that it probably derives from a different neuronal mechanism.

One could claim that these temporal findings do not prove the disappearance of the afterimage mechanism, or even the classical afterimage due to the recent findings of van Lier and colleagues (2009), the filling-in afterimage effect. They showed that there is a very clear negative afterimage effect with stimuli that have inducing chromatic zones presented for only 500 ms (similar to the current effect). This negative afterimage effect is perceived very vividly only when the object contours are maintained after the disappearance of the inducing image (Daw, 1962).

The contour that is presented after the induction has a crucial role in impoundment of the perceived color, assisting in preventing fading of the color into the space (van Lier et al., 2009). Van Lier and colleagues demonstrated the significance of the contours in the filling in effect by using a star-like stimulus comprised of red- and cyan-colored spikes attached to a grey central area (Fig. IX.109-1 in van Lier et al., 2009). They found that the spikes afterimage (with perceived complementary colors) diffused into the grey central area that was not colored before.

The location of the contours in relation to the location of the inducing area can determine which type of effect will be perceived (a positive vs. negative afterimage). In the case that after the disappearance of chromatic inducers the contour of the whole inducing annular is left, or even only the outer contour is left, we observe the complementary color to the chromatic inducer, as has been demonstrated in van Lier et al.'s (2009) illusion. In the case that the inner contour of the inducer is left, after the adaptation, we observe the color dove illusion (positive aftereffect). Since these two effects are elicited in similar spatial and temporal configurations and are contour dependent, it may be suggested that these two effects (color dove and filling-in afterimage) share common underlying mechanisms.

The mechanism that invokes these unique effects should be investigated further. Since the color dove illusion effect (and similarly the filling-in afterimage effect) is rapid and does not require a large area for the inducer region, it is probably performed due to a spatiotemporal edge. We suggest that these effects are actually an aftereffect of the spatiochromatic edge. Analogical to the classic afterimage (where the disappearance of a color that is presented on a large area for a long duration yields the perception of a complementary color), the disappearance of the spatiochromatic edge yields an "afterimage" of a complementary edge. Since the interpretation of the chromatic edge cannot lead to an injective value, the visual system assists with the left contours.

Since the effect involves clear spatial and spectral distinction, it can be derived from an adaptation mechanism of retinal chromatic opponent receptive field (Shevell & Wei, 1998; Spitzer & Barkan, 2005). The adaptation mechanism of the retinal ganglion receptive fields performs a gain control ("curve-shifting") due to their local and remote chromatic characteristics (Shapley & Enroth-Cugell, 1984; Spitzer & Barkan, 2005; Spitzer & Semo, 2002). The aim of this adaptation mechanism (as well as other adaptation mechanisms) is to increase the sensitivity of a specific stimulus domain at a relevant specific stimulus range.

It is possible that the color dove illusion is initiated by this retinal edge adaptation. However, the nature of the manifested effect is determined by a filling-in mechanism that depends on the left contours and is probably a cortical effect (Bressan, Mingolla, Spillmann, & Watanabe, 1997; Komatsu, 2006; Redies & Spillman, 1982; Shevell & Wei, 1998).

REFERENCES

Anstis, S., Rogers, B., & Henry, J. (1978). Interactions between simultaneous contrast and coloured afterimages. *Vision Research*, *18*(8), 899–911.

Barkan, Y., & Spitzer, H. (2009). Color dove illusion. Second place: Best illusion of the year contest. http://illusionoftheyear.com/2009/05/color-dove-illusion/

Blackwell, K. T., & Buchsbaum, G. (1988). Quantitative studies of color constancy. *Journal of the Optical Society of America A*, *5*(10), 1772–1780.

Bressan, P., Mingolla, E., Spillmann, L., & Watanabe, T. (1997). Neon color spreading: A review. *Perception*, *26*(11), 1353–1366.

Daw, N. W. (1962). Why after-images are not seen in normal circumstances. *Nature*, *196*, 1143–1145.

Komatsu, H. (2006). The neural mechanisms of perceptual filling-in. *Nature Reviews Neuroscience*, *7*(3), 220–231.

Redies, C., & Spillman, L. (1982). The neon color effect in the Ehrenstein illusion. *Perception*, *10*(6), 667–681.

Semo, S. (2001). Complementary chromatic induction and remote effects are predicted by an opponent-cell model. *Perception*, *30*(ECVP Abstract Suppl.).

Sezikeye, F. X., & Gurnsey, R. (2009). Simultaneous contrast is size dependent but does not scale with eccentricity. *Journal of Vision*, *9*, 352.

Shapiro, A., Lu, Z.-L., Knight, E., & Ennis, R. (2009). The break of the curve ball. Winner: Best illusion of the year contest. http://illusionoftheyear.com/cat/top-10-finalists/2009/#post-1060.b

Shapley, R., & Enroth-Cugell, C. (1984). Visual adaptation and retinal gain controls. *Progress in Retinal Research*, *3*, 263–364.

Shevell, S. K., & Wei, J. (1998). Chromatic induction: Border contrast or adaptation to surrounding light? *Vision Research*, *38*(11), 1561–1566.

Spitzer, H., & Barkan, Y. (2005). Computational adaptation model and its predictions for color induction of first and second orders. *Vision Research*, *45*(27), 3323–3342.

Spitzer, H., & Semo, S. (2002). Color constancy: A biological model and its application for still and video images. *Pattern Recognition*, *35*(8), 1645–1659.

Tsofe, A. (2004). The role of retinal remote adaptation for perceived "real" colors and subjective colors (Master's thesis). Tel Aviv University, Tel Aviv, Israel.

van Lier, R., Vergeer, M., & Anstis, S. (2009). Filling-in afterimage colors between the lines. *Current Biology*, *19*(8), R323–R324.

Zaidi, Q., Yoshimi, B., Flanigan, N., & Canova, A. (1992). Lateral interactions within color mechanisms in simultaneous induced contrast. *Vision Research*, *32*(9), 1695–1707.

Chapter 110
Blur Adaptation and Induction

Michael A. Webster

The image on the retina is always blurred because of optical aberrations of the eye. Yet typically the world does not "look" blurred, even for individuals with large refractive errors. Moreover, the acuity of the eye varies dramatically from the center of gaze to the periphery, but the outside world generally "feels" focused throughout the visual field. This perception of focus is one of many illusions where the brain appears unaware of its own imperfections, or, more correctly, where it compensates for these imperfections in order to tie percepts to the world rather than to the vagaries of an individual's sensitivity limits (Breitmeyer, 2010). In the case of blur, this compensation may include calibrating spatial coding by adjusting visual sensitivity so that it is matched to the spatial structure of the environment (Webster, 2011, 2015). When the level of blur in the image changes, the perception of focus rapidly readjusts—or adapts—to the change (Webster, Georgeson, & Webster, 2002).

Figure IX.110-1 illustrates adaptation to blur. To experience the effect one should first fixate the spot between the top two images for several seconds and then switch gaze to the spot between the lower pair, The lower pair are both physically focused, but the image on the right (or left) may appear blurrier (or sharper) after adapting, because it falls on a part of the retina that was previously adapted to a physically sharpened (or blurred) stimulus. An animation demonstrating this effect is provided on the Companion Website. The aftereffects are similar for a diverse range of images, from natural to noise, and also show strong transfer across images (Webster et al., 2002). Thus the adaptation is largely to the attribute of blur, irrespective of the specific image carrying it. These effects show that the calibration of subjective focus depends strongly on the immediate preceding context. Aftereffects of blur adaptation have also been found by using measures of visual acuity. Observers wearing lenses that blur their vision adapt to the defocus, and after an hour or more their acuity with the lenses on improves (Mon-Williams, Tresilian, Strang, Kochhar, & Wann, 1998; Pesudovs & Brennan, 1993). Extremely persistent aftereffects of blur have also been reported when observers are exposed to blur for long periods. In an individual whose congenital cataracts were removed in midlife, square-wave edges appeared too sharp even when tested more than a year after the surgery (Fine, Smallman, Doyle, & MacLeod, 2002).

Perceived focus also depends strongly on the spatial context (Webster et al., 2002). For example, the middle bars in Figure IX.110-2 all have square-wave edges, yet when the abutting bars are blurred the middle edges appear sharpened or vice versa. This induction does not depend on precisely aligned edges, for a similar effect can be seen in Figure IX.110-3, where the physically focused center images appear different because the surrounding images have again been blurred or sharpened. These adaptation and induction effects occur not only for spatial blur but also for images blurred in time (Bilson, Mizokami, & Webster, 2005). For instance, a step change in time from bright to dark appears sluggish after adapting to a sharpened temporal change. Blur adaptation also occurs when the edges or images vary in color rather than luminance, though the blur induction may be absent in pure color patterns (Webster, Mizokami, Svec, & Elliott, 2006), even though there is good sensitivity to blur in red-green edges (Wuerger, Owens, & Westland, 2001).

Contextual effects are of course not unique to blur but occur for most stimulus dimensions, for there are countless examples of adaptation aftereffects and simultaneous contrast (Schwartz, Hsu, & Dayan, 2007; Webster, 2011, 2015). Moreover, the effects for blur are reminiscent of the classic illusions of size or spatial frequency adaptation and induction. For example, after adapting to bars with a particular spatial frequency, gratings with lower or higher frequencies will appear shifted in size away from the adapting frequency (Blakemore & Sutton, 1969). It remains unclear to what extent the two adaptation effects reflect different mechanisms, for the neural encoding of blur itself remains unclear (Field & Brady, 1997; Watson & Ahumada, 2011). Yet there are intriguing differences between how adaptation affects single spatial frequencies versus the ratio of spatial frequencies to which blur corresponds. One is that adapting to a physically focused square wave, like the middle bars in Figure IX.110-2, does not produce a blur aftereffect, presumably because it is already in focus and thus does not bias this ratio. Yet the same adapting stimulus reduces contrast sensitivity much more at low frequencies than high, when measured with detection thresholds in single test gratings (Nachmias, Sansbury, Vassilev, & Weber, 1973; Tolhurst, 1972; Webster & Miyahara, 1997). Blur aftereffects are also better predicted by the actual

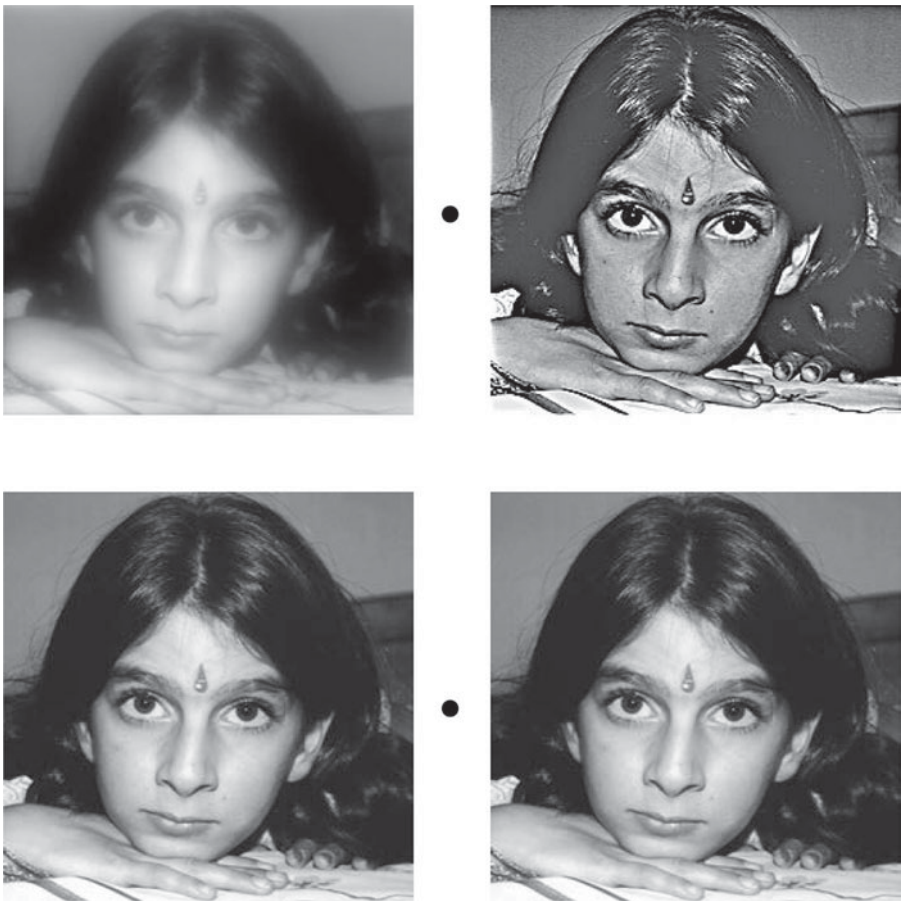


Figure IX.110-1. An example of blur adaptation. Fixate between the top two images for several seconds, and then shift to the spot between the lower images. The image on the right will briefly appear blurrier. (After Webster et al., 2002.)

physical focus of the images than by their amplitude spectra (Webster et al., 2002). The latter can differ among focused images because the spectrum depends not only on the profile of edges but also on the density of edges, so that a focused image with more uniform regions will have a steeper spectrum and thus present less high-frequency contrast to adapt to (Field & Brady, 1997). Yet this image does not behave like a blurred image in the adaptation, suggesting again that the aftereffects cannot be accounted for by the independent sensitivity changes at different frequencies. Such results raise the possibility that the visual system encodes blur as an explicit feature, rather than implicitly by the relative activity across spatial frequency channels.

A second important difference between blur and spatial frequency adaptation is that blur is much more characteristic of the natural, broadband stimuli that the visual system normally encounters and thus may better reveal how spatial vision is calibrated through adaptation (just as much



Figure IX.110-2. An example of blur induction. The middle row of bars are all identical square-wave edges but look sharpened when surrounded by blurred bars (left) or blurred when surrounded by sharpened bars (right). (After Webster et al., 2002.)



Figure IX.110-3. Blur induction in complex images. The two center faces are both focused, but the image on the right appears blurrier because the surrounding faces are sharpened. The effect is best seen in the periphery, by fixating the dot between the two arrays. (After Webster et al., 2002.)

has been revealed about color perception and adaptation by examining natural broadband color signals rather than single wavelengths; Foster, 2011). For instance, adapting to a single spatial frequency biases the perceived size of other gratings but does not affect the apparent size of the adapting grating itself, a pattern known as repulsion aftereffects. Adaptation to blur instead produces a renormalization aftereffect, where the adapting stimulus itself looks more neutral or focused with adaptation. That is, blurred images look less blurred and sharp images appear less sharp as one continues to view them (Elliott, Georgeson, & Webster, 2011). This pattern suggests that the point of subjective focus represents a norm in spatial coding, and the level of blur or sharpness may be represented by the deviation from this norm, similar to the way that colors are coded by how they deviate from the norm of gray (Webster, 2011, 2015). This is counterintuitive from an optical perspective, for a lens can blur an image but never oversharpen it. Yet at the neural level, spatial vision could be off by giving too much weight to either lower or higher frequencies, and again the calibration of image focus depends on balancing this ratio. Finally, it should be noted that this balance does not correspond to a physically balanced amplitude spectrum in the stimulus. Natural images have more energy at lower frequencies (and amplitude falls roughly as the inverse of frequency, or 1/f; Field, 1987). Vision may have adapted over evolutionary timescales to this characteristic spatial structure of images but must still be fine-tuned by processes like sensory adaptation to calibrate individuals and recalibrate them as they develop or age (or get a new pair of glasses). Because "blur" in the environment is largely constant, most of the variation in retinal image blur is due to differences in optical quality between observers. Yet there are environments or contexts that are themselves blurry and that could induce a change in adaptation. One example is low-visibility environments like fog or underwater (Ross, 1974). Another interesting case is when observers must work with images that are blurry. Radiologists can spend hours looking at medical images that are often significantly more blurred than natural images. These images could induce strong blur aftereffects and thus might alter the adapted state of the radiologist (Kompaniez, Abbey, Boone, & Webster, 2011, 2015).

The images in Figures IX.110-1, IX.110-2, and IX.110-3, were blurred or sharpened by varying the slope of the amplitude spectrum. This is very different from optical blur, and a number of studies have explored how perceived focus adapts to changes in the patterns of blur generated by actual aberrations in the eye. How vision adjusts to this optical blur is important for understanding both the consequences of optical errors and their correction. Blur aftereffects are selective for the axis of astigmatism and thus could potentially "correct" for at least some of the patterns of blur specific to the individual's eyes (Sawides et al., 2010). A recent study also found that this adaptation could become more rapid and complete when observers were re-exposed to the astigmatic blur on successive days, suggesting that the visual system might be capable of learning and storing different adaptation states for different contexts (Yehezkel, Sagi, Sterkin, Belkin, & Polat, 2010). This effect parallels a common report of patients that it is difficult to adjust to the optical distortions introduced by a new refractive correction, yet once adjusted, vision with their glasses seems fine the moment they are worn.

The finding that adaptation can be selective for the axis of astigmatism is perhaps not surprising, because there are again well-known examples of orientation-selective adaptation such as the tilt aftereffect (Gibson & Radner, 1937). More intriguingly, Sawides et al. (2010) found that the astigmatism was inducing adaptation not only to the

blur but also to how this oriented blur was altering the perceived shape of the images. For example, Figure IX.110-4 shows a face blurred by positive or negative astigmatism. The images not only appeared blurred but look like different individuals, and adapting to them produces not only a blur aftereffect but also a configural or face aftereffect. This suggests that how the visual system adapts to particular patterns of blur, or other stimulus attributes, may be complex and arise at multiple levels. Adaptation to attributes like spatial frequency or orientation is thought to first arise in visual cortex because of the pattern selectivity and interocular transfer of the aftereffects (Graham, 1989). But if a change in the stimulus induces a change in other perceptual attributes, like the perceived identity of a face, then these differences may also trigger adaptation in the visual processes encoding these attributes.

How adaptation adjusts to astigmatism relates to a well-known cautionary tale in vision science known as the El Greco fallacy (Anstis, 2002). El Greco painted figures that were vertically stretched, and a frequent suggestion was that this might have been because he had astigmatism and thus saw the world as stretched. The fallacy is that this would have distorted his perception of both the painting and the person he was trying to depict and thus could not account for why he would render them differently. However, there is a potential fallacy to the fallacy—if the person and painting subtended different sizes in the retinal image, then the blur could affect them differently. The reason is because the blur is constant with visual angle and would thus produce a larger distortion in a face with a smaller angle. This is illustrated in the larger images in Figure IX.110-4, which show two ways of magnifying the smaller image at top left. In the middle, the image was blurred and then magnified, which preserved the shape in the smaller and larger images but amplified the blur. On the right, the focused face was magnified and then blurred. In this case the level of blur is preserved, but the shapes differ because the blur distorts the larger face less. This is not an argument for whether El Greco was astigmatic or for why he painted stretched images. Rather, the point is that if observers have astigmatism, then, unless they are adapted to discount it, they should experience failures of shape constancy as objects become closer or farther from them. In fact, observers do exhibit long-term adaptation to their uncorrected native astigmatism (Vinas, Sawides, De Gracia, & Marcos, 2012; Marcos et al. 2015).

Defocus and astigmatism are low-order aberrations or simpler optical errors that are typically corrected by lenses or contacts. The eye also includes high-order aberrations that are usually not corrected, so that they may present more chronic patterns of blur. Several studies have explored how the visual system adapts to these high-order aberrations by taking advantage of adaptive optics displays to fully correct the aberrations or simulate new aberrations. In the initial study of these effects, Artal et al. (2004) showed that observers prefer their own aberration pattern to rotated versions, implying that they are naturally adapted to their own intrinsic blur. Observers also consistently perceive their own magnitude of blur as better focused than blur with different magnitudes drawn from their own or others' eyes (Sawides, de Gracia, Dorronsoro, Webster, & Marcos, 2011a, 2011b). Importantly, this is not simply a learned convention or criterion for judging the blur but reflects an actual adaptation to the level of blur to which we are normally exposed. For example, in two observers who differ in their optical quality, both will show an aftereffect when exposed to the blur level from the other person's eye but little aftereffect from their own eye (Sawides et al., 2011a). Thus the perceptual norm for image focus probably does correspond to an actual neural norm within the mechanisms encoding blur. The two eyes of a single observer also often differ in the degree of optical



Figure IX.110-4. Blur adaptation and astigmatism. The two images on the left are blurred by simulating horizontal (top) or vertical (bottom) axes of astigmatism, which causes them to look like pictures of different individuals. The middle image is a magnified version of the top left image and has the same shape but more blur. The right image was blurred after magnifying and has the same amount of blur as the top left image but differs in its relative spatial profile. (After Sawides et al., 2010.)

blur, but the perception of focus through either eye corresponds closely to the blur level of the eye with better optical quality. This suggests that the norm is established at a completely binocular site (Radhakrishnan et al. 2015). Again, such adaptations help contribute to the illusion that our vision is largely error free, because these errors have been discounted from our perception.

REFERENCES

Anstis, S. M. (2002). Was El Greco astigmatic? *Leonardo*, *35*, 208.

Artal, P., Chen, L., Fernandez, E. J., Singer, B., Manzanera, S., & Williams, D. R. (2004). Neural compensation for the eye's optical aberrations. *Journal of Vision*, *4*(4), 281–287.

Bilson, A. C., Mizokami, Y., & Webster, M. A. (2005). Visual adjustments to temporal blur. *Journal of the Optical Society of America A*, *22*(10), 2281–2288.

Blakemore, C., & Sutton, P. (1969). Size adaptation: A new aftereffect. *Science*, *166*(902), 245–247.

Breitmeyer, B. (2010). *Blindspots: The many ways we cannot see*. Oxford: Oxford University Press.

Elliott, S. L., Georgeson, M. A., & Webster, M. A. (2011). Response normalization and blur adaptation: Data and multi-scale model. *Journal of Vision*, *11*(2), 7.

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, *4*(12), 2379–2394.

Field, D. J., & Brady, N. (1997). Visual sensitivity, blur and the sources of variability in the amplitude spectra of natural scenes. *Vision Research*, *37*(23), 3367–3383.

Fine, I., Smallman, H. S., Doyle, P., & MacLeod, D. I. (2002). Visual function before and after the removal of bilateral congenital cataracts in adulthood. *Vision Research*, *42*(2), 191–210.

Foster, D. H. (2011). Color constancy. *Vision Research*, *51*(7), 674–700.

Gibson, J. J., & Radner, M. (1937). Adaptation, after-effect and contrast in the perception of tilted lines: I. Quantitative studies. *Journal of Experimental Psychology*, *20*, 453–467.

Graham, N. V. (1989). *Visual pattern analyzers*. Oxford: Oxford University Press.

Kompaniez, E., Abbey, C. K., Boone, J. M., & Webster, M. A. (2011). Adaptation and the perception of structure in radiological images [Abstract]. *Jounral of Vision*, *11*(1), 12.

Kompaniez, E., Abbey, C. K., Boone, J. M., & Webster, M. A. (2015). Adaptation and visual search in mammographic images. Attention, Perception, and Psychophysics *77*(4), 1081–1087.

Marcos, S., Velasco-Ocana, M., Dorronsoro, C., Sawides, L., Hernandez, M., & Marin, G. (2015). Impact of astigmatism and high-order aberrations on subjective best focus. *Journal of vision*, *15*(11), 4.

Mon-Williams, M., Tresilian, J. R., Strang, N. C., Kochhar, P., & Wann, J. P. (1998). Improving vision: Neural compensation for optical defocus. *Proceedings of the Royal Society B*, *265*(1390), 71–77.

Nachmias, J., Sansbury, R., Vassilev, A., & Weber, A. (1973). Adaptation to square-wave gratings - in search of the illusive third-harmonic. *Vision Research*, *13*(7), 1335–1342.

Pesudovs, K., & Brennan, N. A. (1993). Decreased uncorrected vision after a period of distance fixation with spectacle wear. *Optometry and Vision Science*, *70*(7), 528–531.

Radhakrishnan, A., Dorronsoro, C., Sawides, L., Webster, M. A., & Marcos, S. (2015). A cyclopean neural mechanism compensating for optical differences between the eyes. *Current Biology*, *25*(5), R188–R189.

Ross, H. (1974). *Behavior and perception in strange environments*. London: Allen & Unwin.

Sawides, L., de Gracia, P., Dorronsoro, C., Webster, M., & Marcos, S. (2011a). Adapting to blur produced by ocular high-order aberrations. *Journal of Vision*, *11*(7), 21.

Sawides, L., de Gracia, P., Dorronsoro, C., Webster, M. A., & Marcos, S. (2011b). Vision is adapted to the natural level of blur present in the retinal image. *PLoS ONE*, *6*(11), e27031.

Sawides, L., Marcos, S., Ravikumar, S., Thibos, L., Bradley, A., & Webster, M. (2010). Adaptation to astigmatic blur. *Journal of Vision*, *10*(12), 22.

Schwartz, O., Hsu, A., & Dayan, P. (2007). Space and time in visual context. *Nature Reviews Neuroscience*, *8*(7), 522–535.

Tolhurst, D. J. (1972). Adaptation to square-wave gratings: Inhibition between spatial frequency channels in the human visual system. *Journal of Physiology*, *226*(1), 231–248.

Vinas, M., Sawides, L., De Gracia, P., & Marcos, S. (2012). Perceptual adaptation to the correction of natural astigmatism. *PLoS One*, *7*(9).

Watson, A. B., & Ahumada, A. J. (2011). Blur clarified: A review and synthesis of blur discrimination. *Journal of Vision*, *11*(5), 10.

Webster, M. A. (2011). Adaptation and visual coding. *Journal of Vision*, *11*(5):3, 1–23.

Webster, M. A. (2015). Visual adaptation. Annual Review of Vision Science 1, 547–567.

Webster, M. A., Georgeson, M. A., & Webster, S. M. (2002). Neural adjustments to image blur. *Nature Neuroscience*, *5*(9), 839–840.

Webster, M. A., & Miyahara, E. (1997). Contrast adaptation and the spatial structure of natural images. *Journal of the Optical Society of America A*, *14*(9), 2355–2366.

Webster, M. A., Mizokami, Y., Svec, L. A., & Elliott, S. L. (2006). Neural adjustments to chromatic blur. *Spatial Vision*, *19*(2–4), 111–132.

Wuerger, S. M., Owens, H., & Westland, S. (2001). Blur tolerance for luminance and chromatic stimuli. *Journal of the Optical Society of America A*, *18*(6), 1231–1239.

Yehezkel, O., Sagi, D., Sterkin, A., Belkin, M., & Polat, U. (2010). Learning to adapt: Dynamics of readaptation to geometrical distortions. *Vision Research*, *50*(16), 1550–1558.

PART X
Conflicting Scale or Other Information

Chapter 111
Hybrid Image Illusion

Aude Oliva and Philippe G. Schyns

THE PHENOMENON: PERCEIVING TWO FOR THE SIGHT OF ONE

A hybrid image is a single picture that combines the low spatial frequencies (LSF) of one image with the high spatial frequencies (HSF) of another image, producing a new image with an interpretation that changes with duration, viewing distance, or size (Brady & Oliva, 2012; Oliva & Schyns, 1997; Oliva, Torralba, & Schyns, 2007; Schyns & Oliva, 1994, 1999). Figure X.111-1 shows the popularized Marilyn Monroe—Albert Einstein hybrid image, created in 2007 for *New Scientist* magazine. Very close up, we perceive Albert Einstein's portrait, with some seemingly randomly cast shadows. If we step back, squint our eyes, or reduce the size of the hybrid (such as by making the image smaller on the screen), we will gradually start seeing Marilyn Monroe's face: at first we may perceive a blend of the two portraits (such as Marilyn with a mustache!), but if we go far enough, or look at a tiny image of the hybrid, we should see only Marilyn Monroe, with Einstein's face completely gone from perception. The same principle applies in the real world: if we are standing far away from the object we are looking at, then all we can see are the LSF. For example, if we are looking at the face of someone standing a few tens of meters away from us, we can tell if the face is male or female, but we might not be able to clearly distinguish the facial emotion or the age of the person (for a review, see Sowden & Schyns, 2006).

Our visual system decomposes the information of an image into components termed spatial frequencies (from low to high; Robson, 1966). Figure X.111-1 shows image thumbnails of the spatial frequency bands of the hybrid: the lowest spatial frequencies depict only the shape of Monroe's head and bust. Low spatial frequencies are global luminance variations in the image and convey broad contours: they are literally a blurry, out-of-focus version of the original image. When adding medium spatial frequencies, Monroe's eyes, mouth, nose, ears, and hairline start becoming clear. In contrast to the broad contours found in LSF, the HSF of an image represent sharp details and fine contours, such as the wrinkles of a face or the texture of a surface. The HSF bands of an image resemble a highly detailed line drawing. In Figure X.111-1, we can discern Albert Einstein's portrait in the right-hand thumbnails: at these HSF we can see the details of his moustache, the wisps of his hair, and the sharp lines around his tie and collar.

Superimposing the LSF and the HSF components of two different images on each other creates *hybrid images*. In its simplest version, a hybrid overlaps one coarse, blurry picture with one fine, detailed picture. In practice, image-processing software is used to filter one image with a low-pass filter and the second image with a high-pass filter (Oliva, Torralba, & Schyns, 2006).

THE PSYCHOPHYSICS OF HYBRID IMAGES

In the real world, all the spatial frequency bands of an image correspond to a single and coherent percept. Hybrid images change this rule, leading the visual system to switch from one percept to another. Hybrid illusions have been used in various behavioral and psychophysical studies to show that the order in which frequency bands are used in visual perception depends on the task, and it is not necessarily driven by a fixed schedule (e.g., first mandatory perception of the LSF then perception of the HSF).

Dr. Angry or Ms. Smile?

When we categorize a particular face as old or young, our visual system automatically knows the scale at which to find useful information. For example, an older face will tend to possess many fine-grained wrinkles that are typically represented in the HSF of the image (see the wrinkles of Einstein in Figure X.111-1). Which spatial frequencies the visual system uses for different face categorizations (e.g., identity or facial expressions) is thus an important question. In Schyns and Oliva (1999), we presented hybrid stimuli overlapping, for instance, an expressive male in LSF and a neutral female in HSF (and vice versa for the assignment of gender and expression to LSF and HSF). We briefly presented each hybrid stimulus on the screen (for 50 ms) and asked observers to categorize its gender, its identity, and whether or not it was expressive. We found that upon briefly seeing the same hybrid stimulus (e.g., a neutral female in LSF overlapped with an angry male in HSF; Figure X.111-2a), observers perceived an expressive face

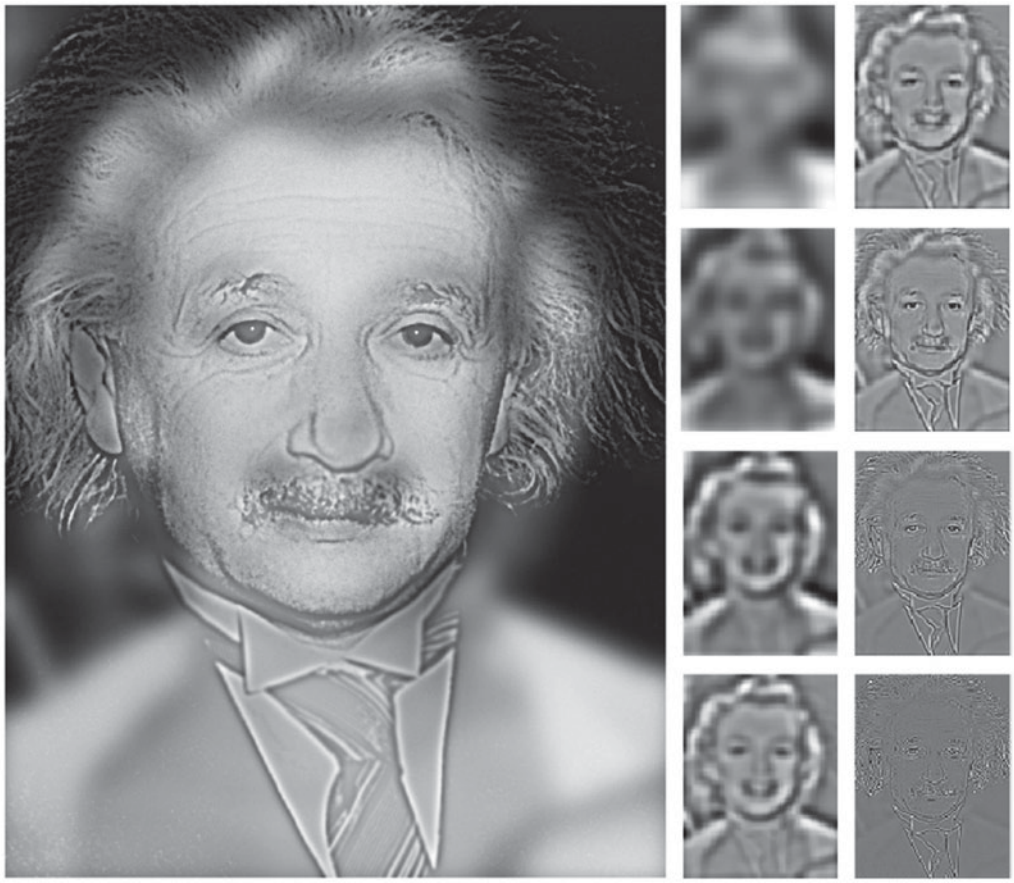


Figure X.111-1. Albert Einstein–Marilyn Monroe hybrid image. The thumbnailed images illustrate the band-pass filtered components of the hybrid, from Monroe (the blurred, or the LSF version) to Einstein (with sharp contours, or the HSF version). (Figure by Aude Oliva.)

(on the basis of HSF) versus "Mary" (on the basis of LSF), depending on the categorization task they were instructed to perform (i.e. "expressive or not?" or "which identity?"). Thus the observers' brains used different spatial frequency information from the same hybrid image to categorize it in different ways, resulting in mutually exclusive perceptions of the same visual input.

Local Jekyll or Global Hyde?

Another important question in vision concerns the local or global nature of the information responsible for everyday recognition. In typical viewing, information useful to identify a face (e.g. the blue color of Paul Newman's eyes) could simply vanish with increasing distance due to the physics of retinal image formation. Identification mechanisms invariant to viewing distances could flexibly adjust to use both local and global cues from the same visual input, depending on which are available on the retina. Here we created an "iHybrid" to study how the visual system uses local and global face identification strategies. An iHybrid represents one identity in LSF and another one in HSF (in Fig. X.111-2b, Brad Pitt and William Macy, respectively). Importantly, the eye fixations of the observer *dynamically* determine the spectral composition of the image (see Fig. X.111-2b). An eye-tracker records in real time the location of the eyes and a computer updates in real time the visual display as a function of the observer's eye position to achieve the following effect: observers see local, full-spectrum information from Brad Pitt in their fovea (i.e., where their eyes fixate, indicated in the figure with a dashed line) and global, complementary information from William Macy outside their parafovea. In Miellet, Caldara, and Schyns (2011), we showed that the visual system indeed identified faces with local and global cues. Consequently, on one trial a given hybrid could be perceived as Brad Pitt by integrating local information over several fixations and on another trial as William H. Macy on the basis of the complementary global information accrued over fewer fixations.

The Asymmetry of Spatial Frequency Integration: Seeing More Than Meets the Eye

Perception is inherently a dynamic process: as we move through the world, the percept of an object we are looking at constantly changes in spatial frequency content (Bonnar, Gosselin, & Schyns, 2002; Pelli, 1999; Sowden & Schyns, 2006), creating an asymmetry in the availability of visual information about the objects we are looking at. When we approach an object or when an object comes toward us, our

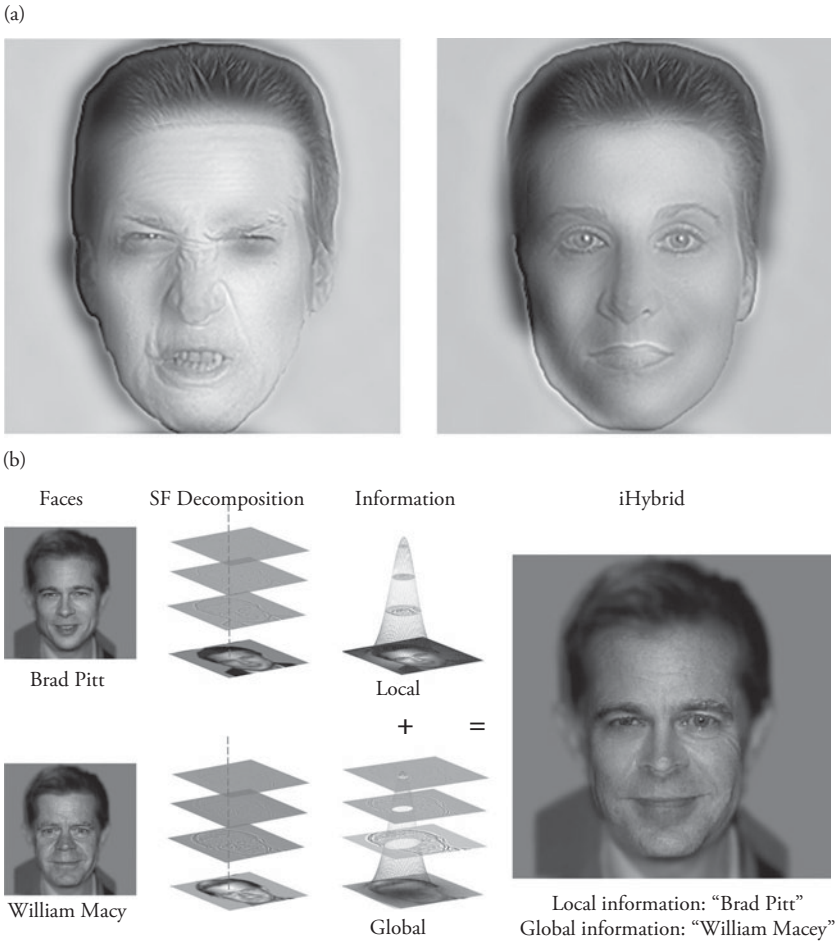


Figure X.111-2. (a) Two hybrid faces used in Schyns and Oliva (1999), each combining a male and a female, with a neutral or expressive face, in either LSF or HSF. (b) Illustration of the iHybrid method of Miellet et al. (2011). An iHybrid simultaneously and dynamically represents two identities across local (here, Brad Pitt) and global (here, William Macy) spectral information, as a function of observer fixation. The local information at fixation (represented with a dotted line) represents a local cut across the spatial frequency spectrum of the first identity, whereas the global information represents the complementary spatial frequency information from the other identity. Decoupling of local and global information in iHybrids enables understanding of the local and global information used for face identification. (Figure by Philippe Schyns.)

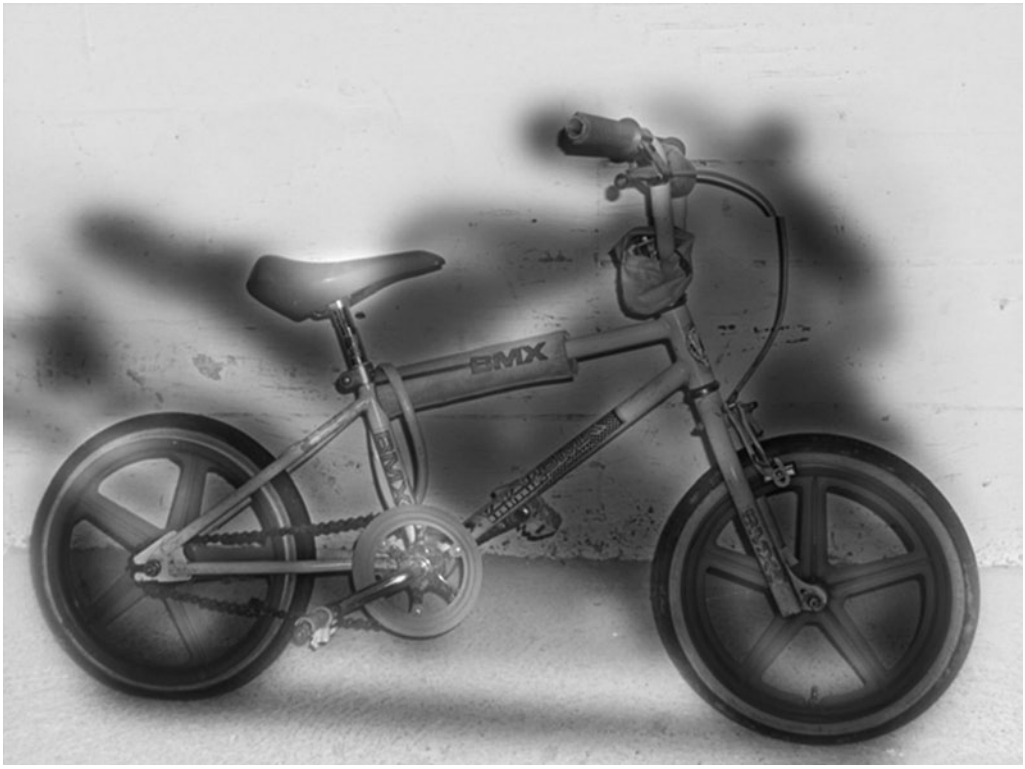


Figure X.111-3. A hybrid image made of a bicycle (seen up-close) and a motorcycle (seen from a distance). Parts of the body of the motorcycle are interpreted as cast shadows up-close, projected on the wall behind the bicycle. (From Oliva et al., 2006.)

perceptual system constantly gains new spatial frequency information about it, adding details to the online representation of the object. But when an object is receding from us, our perceptual system loses information: the object becomes more ambiguous as its details are too high in spatial frequency to be perceived. In Brady and Oliva (2012), using hybrid images we studied how the visual system deals with this asymmetry when integrating spatial frequency information from objects that approach or recede from the observer. Observers judged how similar a hybrid face or hybrid scene was to each of its original images while physically walking toward or away from it or having the stimulus virtually moved toward or away from them on a computer screen. Interestingly, when the observer or the stimulus are approaching each other, observers perceive the stimulus as if the image is simply gaining higher spatial frequency components as they become physically available to the eye. However, when the stimulus or the observer are receding from each other, observers show a *perceptual hysteresis* effect, holding on to fine details that are supposed to be imperceptible at that distance. They can see more HSF information than meets the eyes. This suggests that people naturally make optimal inferences when perceiving objects in the real world by sticking with their previous interpretation when they lose information and constantly reinterpreting their input when gaining new information.

THE ART OF HYBRID IMAGES

Compelling hybrid images can be created with images other than faces: if the mind can reorganize from a distance the shapes of an image into a coherent form, switching between the percepts of two different stimuli can occur. For instance, forms seen as shadows close up will often work to the hybrid's advantage (Oliva et al., 2006). Figure X.111-3 illustrates such an illusion: close up, we see a bicycle with shadows cast on the wall behind. The scene has a three-dimensional flavor to it. But if we step back from the image, what appear to be shadows will regroup to form the body of a motorcycle. At close perceptual range, the parts that made up the motorcycle are reinterpreted as the shadows of the bicycle, localized behind the attended object. This grouping allows for a seamless transformation from one object to another entirely different object.

ACKNOWLEDGMENTS

The writing of this chapter was supported in part by a National Eye Institute grant EY020484 to AO.

REFERENCES

Bonnar, L., Gosselin, F., & Schyns, P. G. (2002). Understanding Dali's *Slave Market with the Disappearing Bust of Voltaire*: A case study in the scale information driving perception. *Perception 31*(6), 683–691.

Brady, T. F., & Oliva, A. (2012). Spatial scale integration during active perception: Perceptual hysteresis when an object recedes. *Frontiers in Perception Science*, *3*, 462.

Miellet, S., Caldara, R., & Schyns, P. G. (2011). Local Jekyll and global Hyde: The dual identity of face identification. *Psychological Science*, *22*(12), 1518–1526.

Oliva, A., & Schyns, P. G. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology*, *34*, 72–107.

Oliva, A., Torralba, A., & Schyns, P. G. (2006). Hybrid images. *ACM Transactions on Graphics*, *25*(3), 527–532.

Pelli, D. G. (1999). Close encounters—An artist shows that size affects shape. *Science 285*, 844–846.

Robson, J. G. (1966). Spatial and temporal contrast sensitivity functions of the visual system. *Journal of the Optical Society of America A*, *56*, 1141–1142.

Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science*, *5*, 195–200.

Schyns, P. G., & Oliva, A. (1999). Dr. Angry and Mr. Smile: When categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition*, *69*, 243–265.

Sowden, P. T., & Schyns, P. G. (2006), Channel surfing in the visual brain. *Trends in Cognitive Sciences*, *10*, 538–545.

Chapter 112
Contrast Asynchronies

Arthur G. Shapiro

INTRODUCTION

A contrast asynchrony (Shapiro, 2008; Shapiro, Charles, & Shear-Heyman, 2005; Shapiro, D'Antona, et al., 2004) is any visual display that produces a temporal conflict between the modulation of first-order information (such as luminance and color) and the asynchronous modulation of second-order information (such as luminance contrast and color contrast). For instance, consider the videos related to Figures X.112-1a and X.112-1b: Video X.112-1a consists of two disks that modulate from light to dark and back again; Video X.112-1b replaces the gray background from Video X.112-1a with a yellow ring that surrounds the left disk and a black ring that surrounds the right disk. The displays in Videos X.112-1a and X.112-1b look radically different from each other: in Video X.112-1a, the disks appear to blink on and off simultaneously, while in Video X.112-1b, the disks appear to alternate (one disk on and the other off, and then they switch).

The videos appear different from each other because there is a conflict between *luminance* and *contrast* information. The luminance information for the disks is the same in both videos (i.e., the disks become light and dark at the same time), but the contrast information is not the same. That is, in Video X.112-1b, the contrasts between the disks and the rings modulate in antiphase: when the disks are white, the contrast between the right disk and black ring is high and the contrast between the left disk and yellow ring is low; when the disks are in their dark phase, the contrast relationships are reversed.

The contrast asynchrony is therefore not a brightness illusion per se but rather a perceptual paradox (Gregory 1997a, 1997b) where observers can see two disks modulating out of phase with each other yet becoming bright and dark at the same time (at 1 Hz). To demonstrate the paradoxical aspect of the display, show Video X.112-1b to a small group of people. All (or nearly all) members of the group will agree that the disks appear to modulate out of phase with each other. Then divide the group into two halves and ask one half to attend to the disk on the right and the other half to attend to the disk on the left; instruct each group to say the word "white" each time the disk that they are attending to becomes white. Everyone in both groups will say the word "white" at the same time, indicating that they see the disks becoming white simultaneously.

The effect of the conflicting luminance and contrast information can also be appreciated in static images. Figure X.112-1c shows five physically identical white disks, each surrounded by different rings ranging from dark on the left to light on the right. There are two observations: the first and most frequently researched is that the appearance of the disks changes from one surround to the next (i.e., each

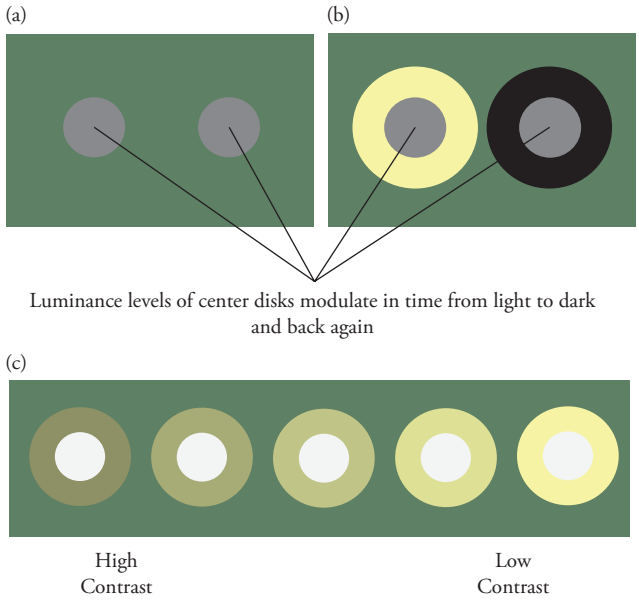
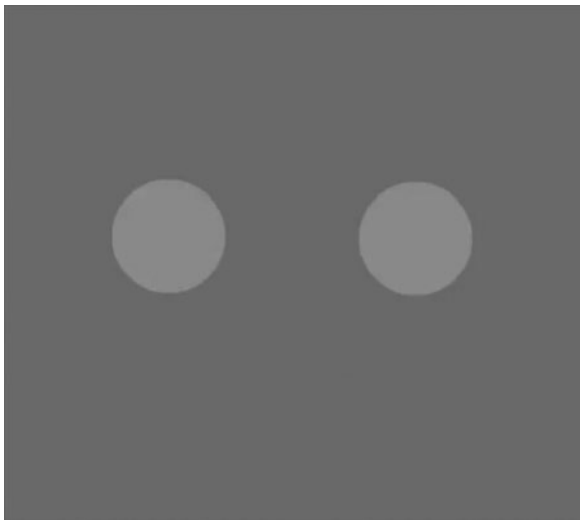
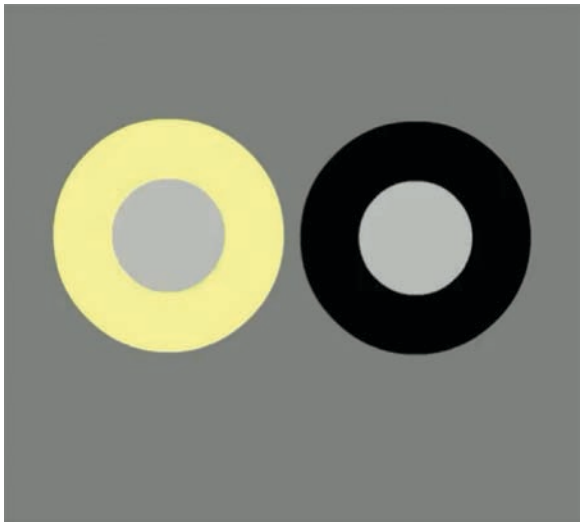


Figure X.112-1. Panels (a) and (b) depict the conditions in Videos X.112-1a and X.112-1b. In both videos, the identical disks modulate in time at 1 Hz from bright to dark to bright, and so on. In Video X.112-1a the disks look like they modulate in phase with each other; in Video X.112-1b the disks paradoxically appear to modulate in antiphase with each other yet get light and dark at the same time. The effect in Video X.112-1b is due to the visual system's ability to separate luminance information (which modulates in phase) and contrast information (the luminance difference between the disks at the surround rings modulates in antiphase). Panel (c) shows five physically identical disks with different surrounds. Most brightness demonstrations concentrate on the different shades of white in the center disks; however, in addition to the appearance of the disk, the visual system also encodes the contrast between the disk and the ring. (Figure by Arthur G. Shapiro.)



Video X.112-1a. See caption for Figure X.112-1.



Video X.112-1b. See caption for Figure X.112-1.

disk could be referred to as white, but the surrounds make the disks appear different in shading); the second observation, which is equally obvious but not researched nearly as frequently, is that the disks change contrast relative to their surround (i.e., the disk on the left has high contrast relative to the dark ring, while the disk on the right has low contrast relative to the light ring). One might suppose that the change in contrast appearance is due to color appearance; that is, the visual system calculates the color of the disk and the color of the ring and then compares the colors—resulting in something like the response, "This disk and ring are similar in brightness; therefore, there is low contrast; and this other disk and ring are not similar in brightness; therefore, there is high contrast."

While people can undoubtedly respond to contrast based on perceived color, an examination of the contrast asynchrony shows that there is also a separate perceptual response based on direct encoding of physical contrast—that is, there exists a response to contrast that does not depend on the identification of colors. Shapiro (2008) suggested that a more accurate model of appearance would contain two separable pathways: one corresponding to contrast and another that generates a response more akin to what we typically think of as color (i.e., hue, brightness, and saturation); the contrast pathway seems to be faster than the "color" pathway. To account for these temporal differences, Shapiro speculated that the asynchrony corresponds to signals encoded early in retinal processing, and the processes that encode the synchronous color modulation might depend on computations the brain performs on the contrast signals. The model Shapiro proposed is similar to other approaches to color and boundary extraction (Grossberg & Todorović, 1988) and may give some indication of the function of different types of color responses found in the visual cortex (e.g., Johnson, Hawken, & Shapley, 2001).

This chapter illustrates three curious aspects of the contrast asynchrony: thin edges can eliminate the asynchronous appearance (Shapiro & Leaver, 2010)*;* contrast responses can affect perceptual organization (Shapiro & Hamburger, 2007, 2011); and contrast asynchronies that occur within a single field are perceptually different from those that occur across two fields (Flynn & Shapiro, 2013; Shapiro et al., 2005).

THIN EDGES CAN ELIMINATE THE ASYNCHRONOUS APPEARANCE

A curious aspect of the contrast asynchrony is that the asynchronous appearance can be completely eliminated without affecting the shading of a central rectangle (Shapiro & Leaver, 2010; Shapiro, Shear-Heyman, et al., 2004). Figure X.112-2a and Video X.112-2a show two identical rectangles that modulate from light to dark and back again; the left rectangle is placed on a black background, and the right is placed on a white background. So, as in Video X.112-1b, the rectangles appear to modulate in antiphase.

However, the same rectangles appear to modulate in phase with each other when thin or thick flanking rectangles (flankers) are placed above and below the modulating rectangles, as in Figure X.112-2b and X.112-2c. The

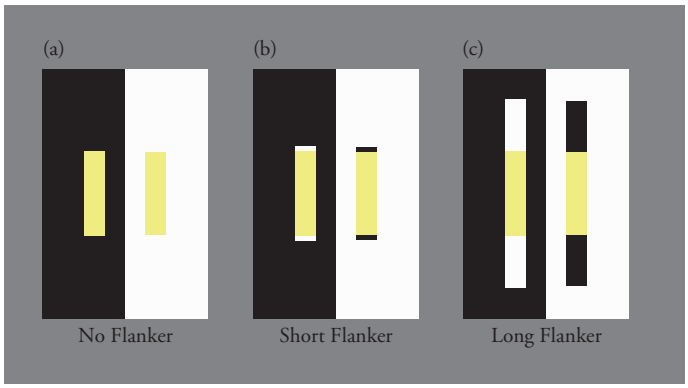
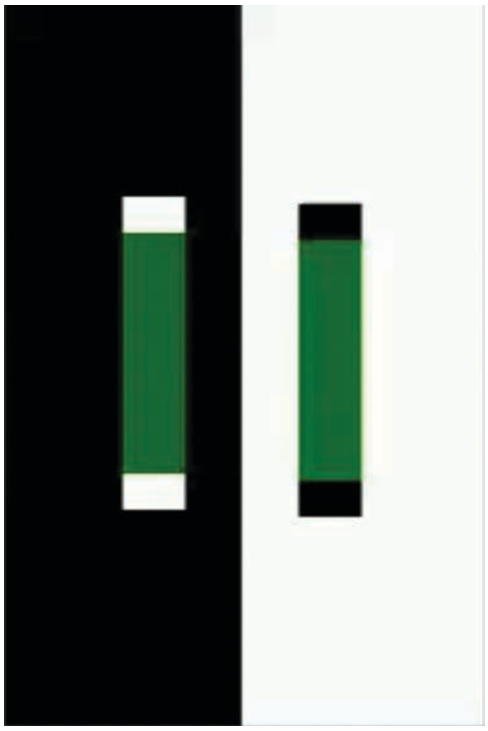


Figure X.112-2. Panels (a), (b), and (c) depict the conditions in Videos X.112-2a, X.112-2b, and X.112-2c. Introducing edges on the top and bottom of the modulating bar can eliminate the perception of asynchrony. In all figures, the yellow rectangles repeatedly modulate in time between bright yellow to dark blue. In panel (a), the rectangles appear to modulate in antiphase (just like in Video X.112-1b). In (b) and (c), flanking edges (flankers) are placed above and below the rectangles; the presence of the flankers makes the rectangles appear as if they are modulating in phase with each other. (Figure by Arthur G. Shapiro.)

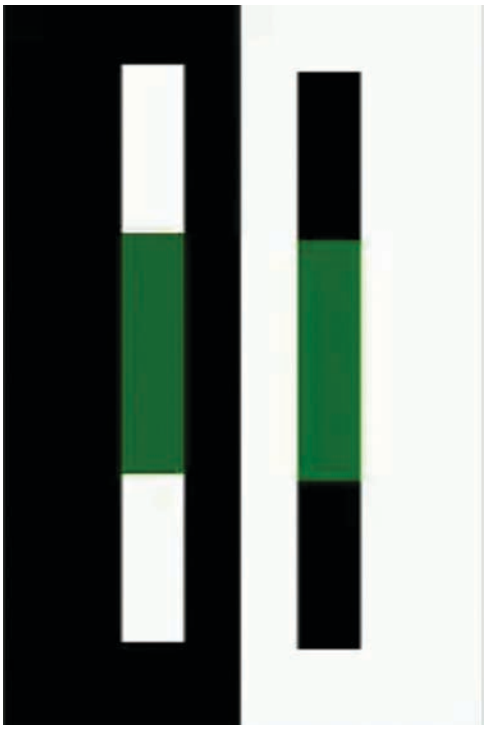


Video X.112-2a. See caption for Figure X.112-2.



Video X.112-2b. See caption for Figure X.112-2.

change of appearance can be seen by viewing Video X.112-2b (in which the flankers are thin, as in Figure X.112-2b) and Video X.112-2c (in which the flankers are thick, as in Figure X.112-2c). The brightness of the flankers is the opposite of the background: the left flanker is bright and the right flanker is dark so that the configuration resembles White's effect.



Video X.112-2c. See caption for Figure X.112-2.

There are two remarkable aspects of the display. First, Shapiro and Leaver (2010) showed that edges as thin as 1 min of visual angle eliminate the asynchrony. So, if an observer were to view video X.112-2b from a distance, the rectangles would appear synchronous until the flanking edges were no longer visible. This perception can also be created by squinting, which will blur the image and thus make the edges less visible.

Second, Shapiro and Leaver (2010) showed that the same edges that eliminate the asynchrony did not have any effect on the appearance of the rectangles when measured in a brightness-matching task. Hence, the perception of modulation can switch from in phase to out of phase even though the brightness of the rectangles does not change at all. The synchronous or asynchronous appearance is related to the contrast between the rectangles and the surrounds, not (as in standard brightness illusions) to the apparent lightness levels of the rectangles induced by the surrounds.

CONTRAST RESPONSES CAN AFFECT PERCEPTUAL ORGANIZATION

The contrast asynchrony also raises questions concerning the perceptual organization of features within an image. A central idea within Gestalt psychology is that of figure–ground separation (Rubin, 1915): a "figure" possesses a "thing character" that is spatially cohesive and enclosed (i.e., demarcated by a contour and filling the entire surface inside its border); a "ground" possesses a "material character" that appears to extend behind the figure. Our perception parses the world into figure and ground based on well-known Gestalt grouping factors (symmetry, proximity,

closure, common fate, etc.). Contrast, however, cuts across the figure–ground separation: the contrast information cannot be considered part of either the figure perceptual unit or the ground perceptual unit. It is therefore worth considering how contrast information affects perceptual grouping (Shapiro & Hamburger, 2007).

Shapiro and Hamburger's (2007) original demonstrations of grouping by contrast were expanded on in a 2009 illusion contest entry (Dixon, Shapiro, & Hamburger, 2011) and in a book chapter (Shapiro & Hamburger, 2011). Here I show an effect based on the same principle but in a visual display that has not previously been published. Panel (a) in Figure X.112-3 shows 16 disks against a uniform background. In Video X.112-3a, the disks modulate in time from light to dark and back again. The rows of disks alternate in phase—that is, when the top row and third row become bright, the second row and fourth row become dark. Perceptual grouping is driven by the similarity of the brightness levels of the disks. Not surprisingly, therefore, rows 1 and 3 group together, and rows 2 and 4 group together. In panel (b) and Video X.112-3b, two of the disks in each row are placed on backgrounds that shift from bright to dark while the backgrounds for the other two disks in each row shift from dark to bright. In the context created by the differing backgrounds, the disks organize in response to the contrast information produced between disks and backgrounds.

The change in the contrast response seems to supersede Gestalt principles that are based on the appearance of the disks themselves (principles such as similarity, proximity, and common region). The effect therefore suggests that the organizational principle depends on information that is contextual. The implication is that the visual response to second-order information may serve as a driver for visual grouping (a result that is similar in principle to texture segregation by Sutter, Beck, & Graham, 1989).

CONTRAST ASYNCHRONIES WITHIN A SINGLE FIELD

A contrast asynchrony can also be created with a single field (Shapiro et al., 2005). An example can be seen in the window-shade/rocking-disk illusion shown in Figure X.112-4. The display consists of a single uniform disk whose luminance modulates in time; the disk is surrounded by a ring that is split so that its top half is white and its bottom half is black. The display fits the definition of a contrast asynchrony because the contrast between the disk and the top and bottom portions of the ring alternates: when the disk is bright, the contrast between the disk and the top of the ring is low and the contrast between the disk and the bottom of the ring is high and vice versa when the disk is dark.

There are two different and perceptually separable effects in this configuration; both effects are quite distinct from the two-field contrast asynchrony shown in Figure X.112-1b. The window-shade illusion occurs when the ring is thick (Video X.112-4a): motion appears to spread across the disk, as if a window shade were being pulled up and down across the surface of the disk. Some observers report a three-dimensional effect, as if a butterfly valve is opening and closing; the three-dimensional effect goes away when the display is rotated 90°, although the shading remains. The rocking-disk illusion occurs when the ring is thin: the disk appears to bounce up and down in a manner similar to the "phenomenal phenomena" of Gregory and Heard (1983) or reverse phi (Anstis & Rogers, 1975). Shapiro et al. (2005) showed that the contrast that produces the rocking disk occurs independently of the contrast that produces shading. The contrast that produces rocking occurs only when the ring is thin (the motion disappears when the ring is greater than about 10 min of visual angle), and the contrast that produces

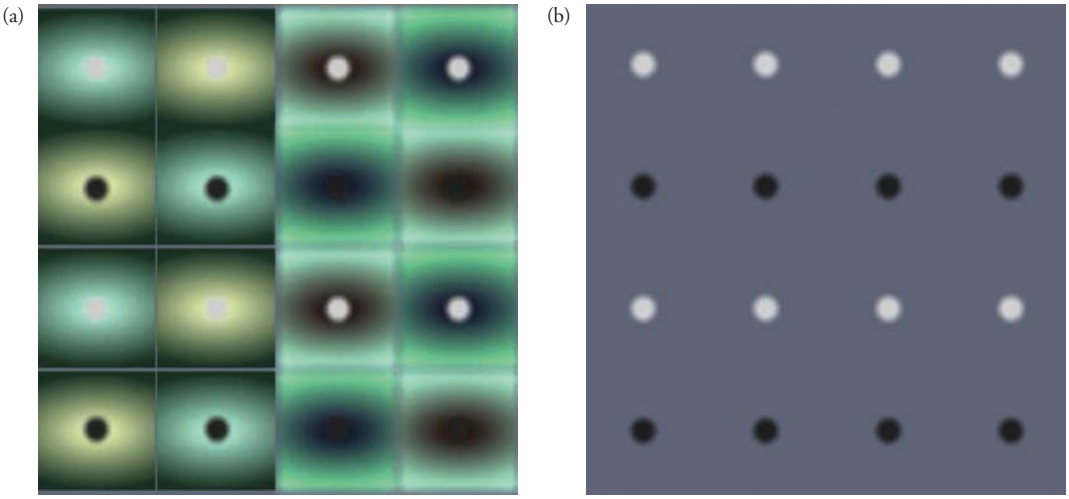
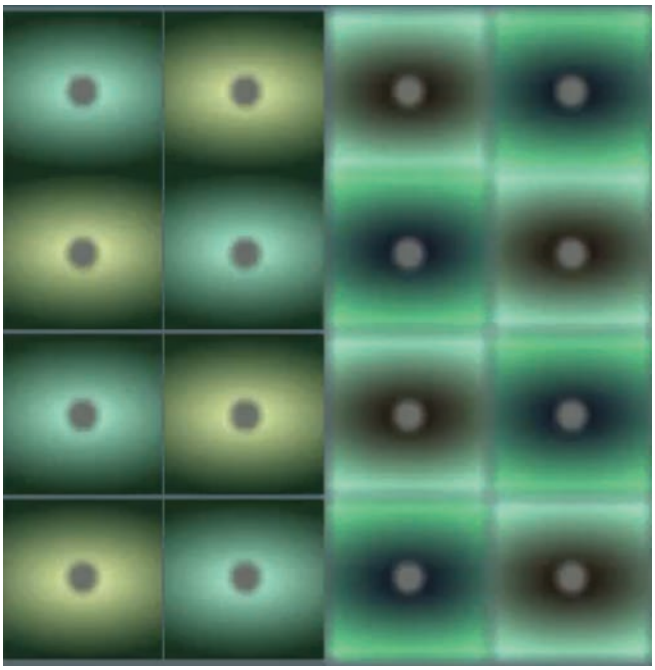


Figure X.112-3. Panels (a) and (b) correspond to Videos X.112-3a and X.112-3b. Both videos contain four rows of disks whose luminance levels modulate from light to dark; rows 1 and 3 modulate in antiphase relative to the dots in rows 2 and 4. In panel (a), bright blurred rings surround the dots on the left and dark blurred rings surround the dots on the right. In panel (b) a uniform gray field surrounds all dots. In Video X.112-3a the dots group with the contrast level—that is, even though the dots in individual rows are identical to each other, the grouping follows a diagonal zig-zag pattern across the display (indicated by the red line); in Video X.112-3b the dots group along the alternate rows. (Figure by Arthur G. Shapiro.)



Video X.112-3a. See caption for Figure X.112-3.

shading becomes stronger as the ring increases in thickness, suggesting that the mechanism(s) responsible for producing shading can integrate over a very large (> 10° of visual angle) area. The window-shade and rocking-disk effects can also occur simultaneously (Video X.112-4b). For instance, if a thin gray line is interposed between the modulating disk and the split ring in the window-shade illusion, the disk will continue to shade but will also appear to bounce up and down.



Video X.112-3b. See caption for Figure X.112-3.

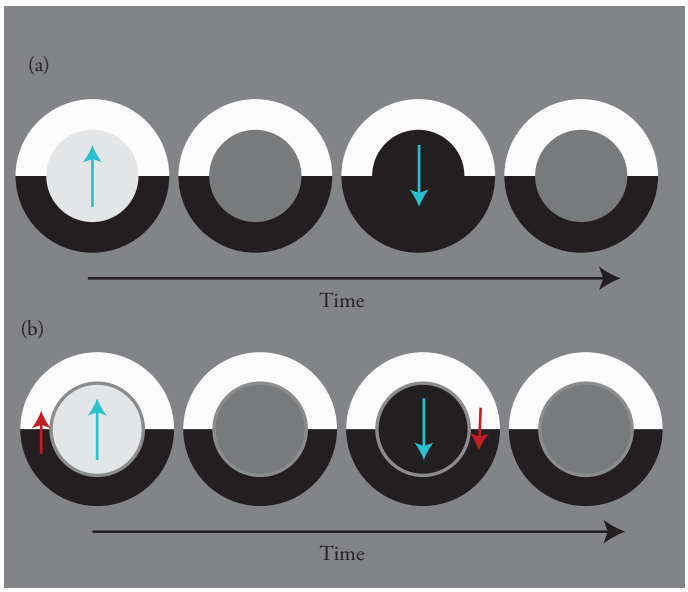
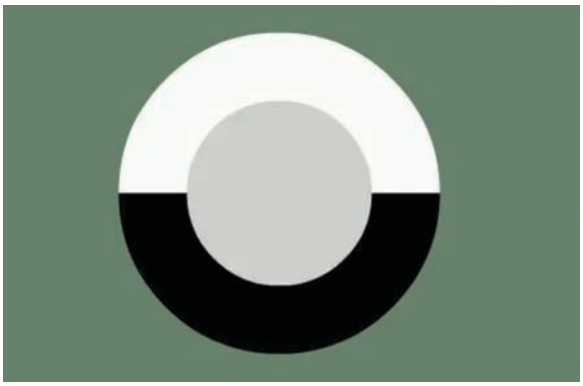
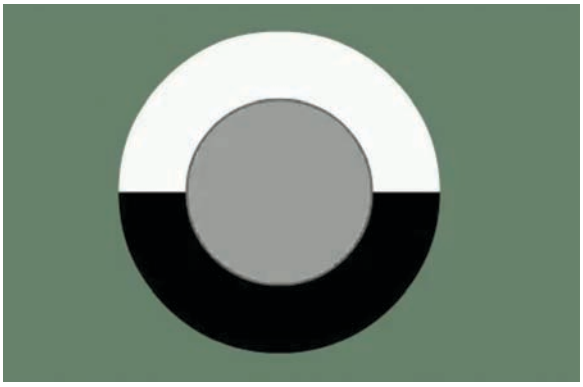


Figure X.112-4. Window-shade/rocking-disk illusions. Panels (a) and (b) correspond to Videos X.112-4a and X.112-4b. In the window shade illusion (panel a), a disk is surrounding by a split ring (half dark, half white). The central disk is spatially uniform, but the disk appears shaded as the luminance of the disk changes from dark to light (almost like a window shade being pulled up and down). The direction of the shading is toward low contrast—that is, as the disk becomes whiter the shade moves toward the white half of the ring, and vice versa for when the disk becomes darker. (Figure by Arthur G. Shapiro.)



Video X.112-4a. See caption for Figure X.112-4.



Video X.112-4b. See caption for Figure X.112-4.

CONCLUSIONS

Contrast asynchronies create a perceptual separation between the contrast information and luminance information and are useful because they separate the responses of visual subsystems. For instance, Shapiro (2008) conducted a systematic investigation of the two-field contrast asynchrony within color space and reported that the asynchrony disappears whenever the color of the disks modulates along one line in a cone-based color space (Derrington, Krauskopf, & Lennie, 1984) and the colors of the rings are fixed at opposite ends of an orthogonal line. So, for example, the disks appear to modulate in phase with each other if the color of the disks changes along an S-cone color line and the color of one surround ring is reddish (i.e., +LM) while the other is greenish (i.e., –LM); the disks appear in antiphase if the colors of the surround rings are at opposite ends of a line closer to the S-cone color line. This finding can be explained by a quantitative model in which there is a separate contrast pathway; the contrast pathway sums the absolute difference between the center disks and surround rings across all color channels.

A model with separate color and color contrast pathways is important because it seems like a possible psychophysical approach for understanding many recent findings in color research. For most of the 20th century, models of color vision consisted of two stages: three linear transducers (the S, M, and L cones) followed by opponent (and additive) combinations of cone responses (Hurvich & Jameson, 1957; Vos & Walraven, 1972a, 1972b). However, such models of color vision provide an insufficient framework for integrating recent physiological findings that suggest "higher-order" color coding (Gegenfurtner & Ennis, 2015); rectified chromatic responses in retinal M ganglion cells (Lee & Sun, 2009); fast color responses in the dorsal occipital lobe and slow color responses in the ventral occipital lobe (Liu & Wandell, 2005); and single-cell recording that indicates that some cells combine color with luminance information (Johnson, Hawken, & Shapley, 2004, 2008) and yet other cells specifically respond to hue (Conway, Hubel, & Livingstone, 2002; Conway, Moeller, & Tsao, 2007; Stoughton & Conway, 2008). Most of these findings seem to confound the difference between color and color contrast, and so while these findings may not be explained by the standard model of color vision, they might be explained by a model in which changes in hue, saturation, or brightness activate both color cells and color contrast cells.

Contrast asynchronies can isolate different uses of contrast information within the visual system. For instance, the different perceptions of the window-shade and rocking-disk effects indicate that the visual system uses the same physical information (i.e., alternation of contrast information) for two perceptual ends: a motion effect and a contrast effect. Furthermore, the effects produced by single-field and two-field asynchronies can be separated by examining dichoptic presentations of the displays. Observers can perceive the window-shade illusion when the split ring is presented to one eye and the modulating disk is presented to the other, but observers cannot see the contrast asynchrony when two modulating disks are presented to one eye and the black and white rings are presented to the other eye (Flynn & Shapiro, 2013). The difference between the dichoptic response and the single-field contrast asynchrony suggests that the standard contrast asynchrony elicits a fast retinal response, whereas the window-shade illusion elicits a slower, cortical contrast response that integrates over a large spatial area; these results are consistent with findings that suggest multiple levels of brightness induction (D'Antona, Christiansen, & Shevell, 2014; Shevell, Holliday, & Whittle, 1992).

There are many aspects of contrast asynchronies that have yet to be investigated thoroughly. The effects probably correspond (in some currently unspecified way) to other functionally based divisions of visual processing. For instance, division between a boundary contour system and feature contour system (Grossberg & Todorović, 1988) and divisions based on object processing versus motion processing almost certainly overlap with categories shown here.

Perhaps the most curious of these perceptual separations involves the interaction between coarse- and fine-scale information. Like hybrid images produced by Oliva and Schyns and hidden images produced by Wade (both in this compendium), contrast asynchronies seem to create different responses at different spatial scales. At the illusion contest in 2007, Shapiro and Knight (2007) showed that blur produces motion not seen in unblurred displays, and thin edges stopped the perception of motion that is physically present at low spatial frequencies. These examples suggest that contrast asynchronies result in paradoxical perceptions that arise when the information in the display places different visual responses in conflict with each other.

REFERENCES

Anstis, S. M., & Rogers, B. J. (1975). Illusory reversal of visual depth and movement during changes of contrast. *Vision Research*, *15*, 957–961.

Conway, B. R., Hubel, D. H., & Livingstone, M. S. (2002). Color contrast in macaque V1. *Cerebral Cortex*, *12*(9), 915–925.

Conway, B. R., Moeller, S., & Tsao, D. Y. (2007). Specialized color modules in macaque extrastriate cortex. *Neuron*, *56*(3), 560–573.

D'Antona, A. D., Christiansen, J. H., & Shevell, S. K. (2014). Separating monocular and binocular neural mechanisms mediating chromatic contextual interactions. *Journal of Vision*, *14*(4), 13.

Derrington, A. M., Krauskopf, J., & Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *The Journal of Physiology*, *357*, 241–265.

Dixon, E. L., Shapiro, A. G., & Hamburger, K. (2011). Grouping by contrast. Second place: Best Illusion of the Year Contest. http://illusionoftheyear.com/2011/grouping-by-contrast/%3E

Flynn, O. J., & Shapiro, A. G. (2013). The separation of monocular and binocular contrast. *Vision Research*, *93*, 19–28.

Gegenfurtner, K. R., & Ennis, R. (2015). Fundamentals of color vision: II. Higher order color processing. In A. J. Elliot, M.D. Fairchild, & Franklin (Eds.), *Handbook of color psychology* (pp. 70–109). Cambridge, UK: Cambridge University Press.

Gregory, R. L. (1997a). Knowledge in perception and illusion. *Philosophical Transactions of the Royal Society B*, *352*(1358), 1121–1127.

Gregory, R. L. (1997b). Visual illusions classified. *Trends in Cognitive Science*, *1*(5), 190–194.

Gregory, R. L., & Heard, P. F. (1983). Visual dissociations of movement, position, and stereo depth: Some phenomenal phenomena. *The Quarterly Journal of Experimental Psychology A*, *35*(Pt. 1), 217–237.

Grossberg, S., & Todorović, D. (1988). Neural dynamics of 1-D and 2-D brightness perception—A unified model of classical and recent phenomena. *Perception & Psychophysics*, *43*(3), 241–277.

Hurvich, L. M., & Jameson, D. (1957). An opponent-process theory of color vision. *Psychological Review*, *64*(6 Pt. 1), 384–404.

Johnson, E. N., Hawken, M. J., & Shapley, R. (2001). The spatial transformation of color in the primary visual cortex of the macaque monkey. *Nature Neuroscience*, *4*(4), 409–416.

Johnson, E. N., Hawken, M. J., & Shapley, R. (2004). Cone inputs in macaque primary visual cortex. *Journal of Neurophysiology*, *91*(6), 2501–2514.

Johnson, E. N., Hawken, M. J., & Shapley, R. (2008). The orientation selectivity of color-responsive neurons in macaque V1. *Journal of Neuroscience*, *28*(32), 8096–8106.

Lee, B. B., & Sun, H. (2009). The chromatic input to cells of the magnocellular pathway of primates. *Journal of Vision*, *9*(2): 15, 1–18.

Liu, J., & Wandell, B. A. (2005). Specializations for chromatic and temporal signals in human visual cortex. *Journal of Neuroscience*, *25*(13), 3459–3468.

Rubin, E. (1915). *Synsoplevede figurer, studier i psykologisk analyse*. Copenhagen: Gyldendal, Nordisk forlag.

Shapiro, A. G. (2008). Separating color from color contrast. *Journal of Vision*, *8*(1).

Shapiro, A. G., Charles, J. P., & Shear-Heyman, M. (2005). Visual illusions based on single-field contrast asynchronies. *Journal of Vision*, *5*(10), 764–782.

Shapiro, A. G., D'Antona, A. D., Charles, J. P., Belano, L. A., Smith, J. B., Shear-Heyman, M. (2004). Induced contrast asynchronies. *Journal of Vision*, *4*(6), 459–468.

Shapiro, A. G., & Hamburger, K. (2007). Grouping by contrast: Figure-ground segregation is not necessarily fundamental. *Perception*, *36*(7), 1104–1107.

Shapiro, A. G., & Hamburger, K. (2011). When is an illusion an illusion? An examination of contrast information in grouping and grid phenomena. In E. P. Charles (Ed.), *A new look at new realism: The psychology and philosophy of E. B. Holt*. (pp. 81–102). Piscataway, NJ: Transaction.

Shapiro, A. G., & Knight, E. (2011). Where has all the motion gone? Third place: Best Illusion of the Year Contest. http://illusionoftheyear.com/2007/where-has-all-the-motion-gone/%3E

Shapiro, A. G., & Leaver, A. M. (2010). Edges can eliminate the appearance of the contrast asynchrony. *Ophthalmic and Physiological Optics*, *30*(5), 534–544.

Shapiro, A. G., Shear-Heyman, M., Milana, M., Charles, J., Leaver, A., Belano, L. A. (2004). Thin edges and the induced contrast asynchrony. *Journal of Vision*, *4*(8), 346a.

Shevell, S. K., Holliday, I., & Whittle, P. (1992). Two separate neural mechanisms of brightness induction. *Vision Research*, *32*(12), 2331–2340.

Stoughton, C. M., & Conway, B. R. (2008). Neural basis for unique hues. *Current Biology*, *18*(16), R698–R699.

Sutter, A., Beck, J., & Graham, N. (1989). Contrast and spatial variables in texture segregation: Testing a simple spatial-frequency channels model. *Perception & Psychophysics*, *46*(4), 312–332.

Vos, J. J., & Walraven, P. L. (1972a). An analytical description of the line element in the zone-fluctuation model of colour vision: I. Basic concepts. *Vision Research*, *12*(8), 1327–1344.

Vos, J. J., & Walraven, P. L. (1972b). An analytical description of the line element in the zone-fluctuation model of colour vision: II. The derivation of the line element. *Vision Research*, *12*(8), 1345–1365.

Chapter 113
Hidden Images

Nicholas J. Wade

It is relatively easy to hide pictorial images, but this is of little value if they remain hidden. The skill is in revealing previously concealed images. This has long been the aim of certain branches of art. *Hidden Images* is the title of a book concerned with visual artifice (Leeman, Elffers, & Schuyt, 1976), but not in the manner examined here. It presents examples of anamorphic art and is an extension of the analysis provided by Baltrusaitis (1976). Anamorphoses have been used to hide images for political and artistic purposes. The former provided a way in which otherwise forbidden messages could be communicated, and the latter displayed mastery of perspective (see Giusti, 2009; Mannoni, Nekes, & Warner, 2004). Presenting hidden images for visual purposes is a more modern preoccupation, and some of the perceptual processes involved in them are described in this chapter.

In the natural world, camouflage has evolved as a form of concealment that has served adaptive purposes of survival against predators or evasion of detection by prey. It has similarly been used in art, but this is not a topic that is considered within this chapter. Pictorial images can be concealed in order to evade detection or to render recognition more difficult. That is, it can be difficult either to detect the presence of a particular component in a picture or to determine the identity of the referent. Thus there are various ways in which images can be concealed or, more precisely, made difficult to detect or recognize. Some of these are described together with graphical examples of their operation, starting with images that are difficult to detect. The initial examples present hidden images of faces. Since pictures of faces are among the stimuli we are most adept at detecting (Bruce & Young, 1998; Calder, Rhodes, Johnson, & Haxby, 2011), it seems surprising that they can be concealed so successfully. Negatives of faces are notoriously difficult to identify (Gilad, Meng, & Sinha, 2009), but there is no problem recognizing that they are representations of faces. In the initial examples shown here, it is difficult to detect that faces are present at all in the illustrations.

Pictorial images can be carried in some pattern or design such that minor variations in the design define facial features that are not initially detected. Figure X.113-1 is an example of what I call "perceptual portraits": these represent people in an unconventional style. The aim of perceptual portraits is both artistic and historical. They generally consist of at least two elements—the portrait and some appropriate motif. The nature of the latter depends on the endeavors for which the portrayed person was known. In some cases, as with Purkinje's portrait in Figure X.113-1, the motif is drawn specifically to display a phenomenon associated with Purkinje: Purkinje (1825) described the distortions that can be seen in patterns of concentric circles (see Figure I.1-13 in chapter 1 of this volume). In others, it is derived from a figure or text in a person's book or an apparatus he or she invented. The portraits and motifs have themselves been manipulated in a variety of ways using graphical, photographical, and computer graphical procedures. The illustrations often require some effort on the part of the viewer to discern the faces embedded in them. It is hoped that such perceptual portraits both attract attention and engage the spectator's interest to a greater degree than do conventional paintings, prints, or photographs. The intention is that the visual intrigue will enhance the viewer's desire to discover why particular motifs have been adopted and in turn to learn more about the persons portrayed: it is intended to be an instance of art serving science. Many more examples can be found in Wade (1990, 1992, 1995, 2011, 2014, 2016) and at http://neuroportraits.eu/.

In order to produce Figure X.113-1, the circles were drawn and then photographed on high-contrast film. Purkinje's portrait was also rendered on high-contrast film so that the main features were black areas on a transparent background. Negatives of the face and design were differentially exposed so that the portrait is defined by minor differences in the thicknesses of the lines (see Wade, 1990, for a more detailed description). The portrait is initially below the threshold for detection, but that threshold is determined by characteristics of the design. High-contrast, high spatial frequency contours can suppress the visibility of low-contrast, low spatial frequency components within them. The portrait can be rendered detectable in a number of ways, all of which involve reducing the high spatial frequency contrast by defocus. The techniques include removing spectacles, if they are worn, or defocusing the pattern by accommodating to a point nearer to or further from the picture plane; viewing the picture from such a distance that the circles are no longer sharply resolved or reducing the size of the image; and moving the head or the image rapidly from side to side. This last procedure produces spokes or propellers radiating from the center, and they appear to

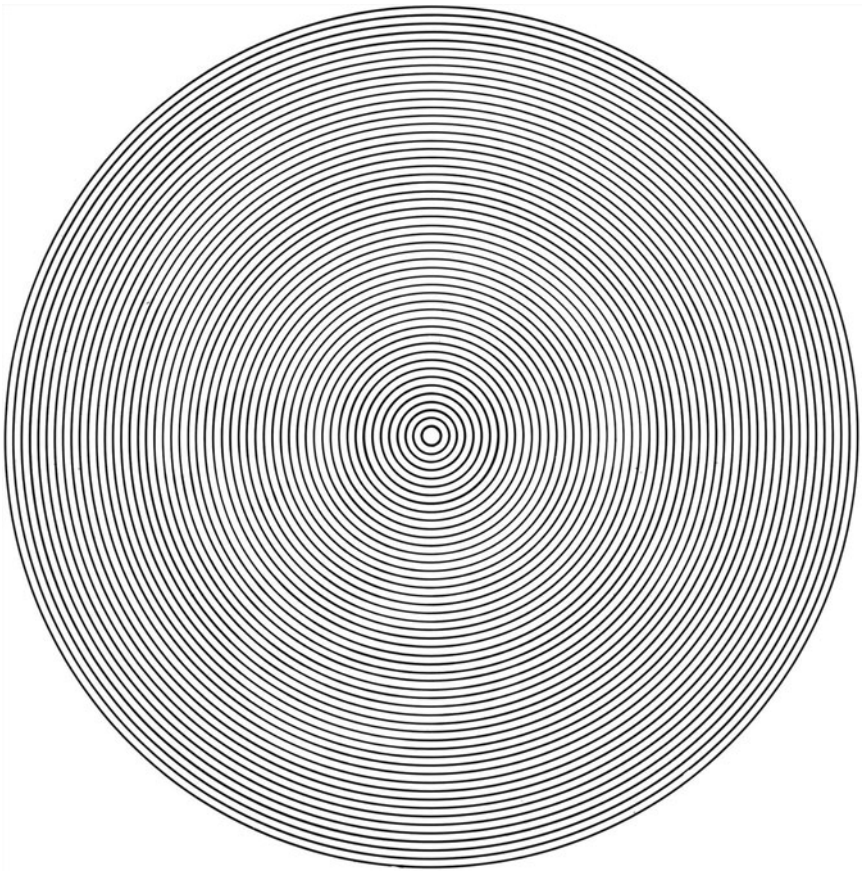


Figure X.113-1. *Sehen in subjectiver Hinsicht* by Nicholas Wade displays a portrait of Jan Evangelista Purkinje. He is shown in three-quarter profile and his left eye is at the center of the design. Purkinje's face is more easily seen if the picture is viewed from beyond reading distance.

rotate if the pattern is moved in a circular path. Similar distortions can be seen with prolonged observation of the pattern when it is stationary, and it is these that Purkinje observed and tried to explain.

The motif in the case of Figure X.113-1 is a pattern like one Purkinje discussed with regard to the distortions visible with it. Variants of illusions can also be the carrier patterns, as in Figure X.113-2, which has features in common with the Zöllner and Poggendorff illusions. The thin vertical lines appear tilted anticlockwise, and they also do not appear to be aligned. What is not immediately apparent is the portrait carried within the design. The bespectacled face of Ian Howard can be seen if the procedures described previously are followed. Once the face has been detected, it will not prove so difficult to see it subsequently. In part this is due to appreciating the spatial scale of the portrait with respect to the design that carries it, as this in turn can determine where to direct the gaze.

In Figures X.113-1 and X.113-2, the patterns carrying the portraits were drawn specifically for their subjects. Any regular and repetitive pattern can be employed as the carrier, and text proves particularly suitable for this purpose because its content and meaning can also speak to the endeavors of the partially portrayed person. Four such perceptual portraits are shown in Figure X.113-3. Two contain portraits of philosophers and the words that carry them are closely associated with their philosophical positions; a third presents a mathematician, and the fourth is of an artist. First, Descartes is combined with a visual variant of his dictum *cogito ergo sum*, which reads "I am seen therefore I am." Descartes' (1637) conviction of the

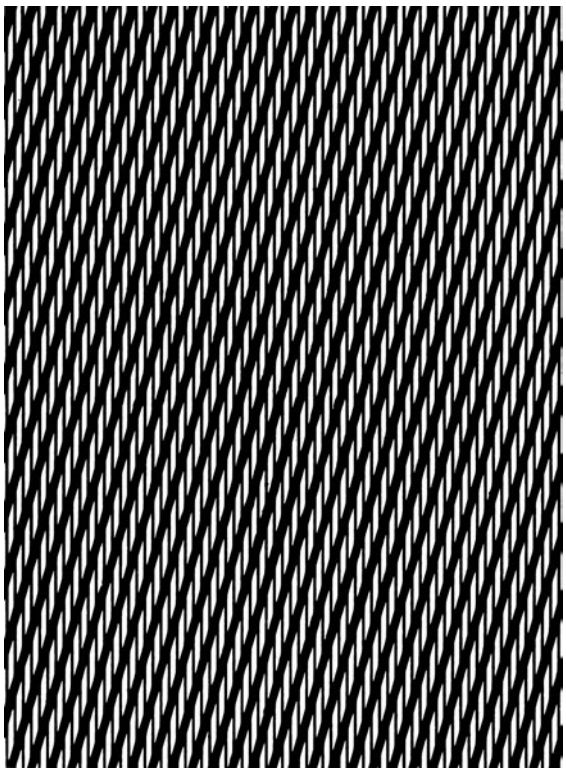


Figure X.113-2. *Human facial orientation I* by Nicholas Wade. The orientations of the vertical lines appear tilted and misaligned. It also contains a portrait of Ian Howard, author of *Human Spatial Orientation* (1966) and *Human Visual Orientation* (1982). The design is the cover of the dust jacket of the latter.

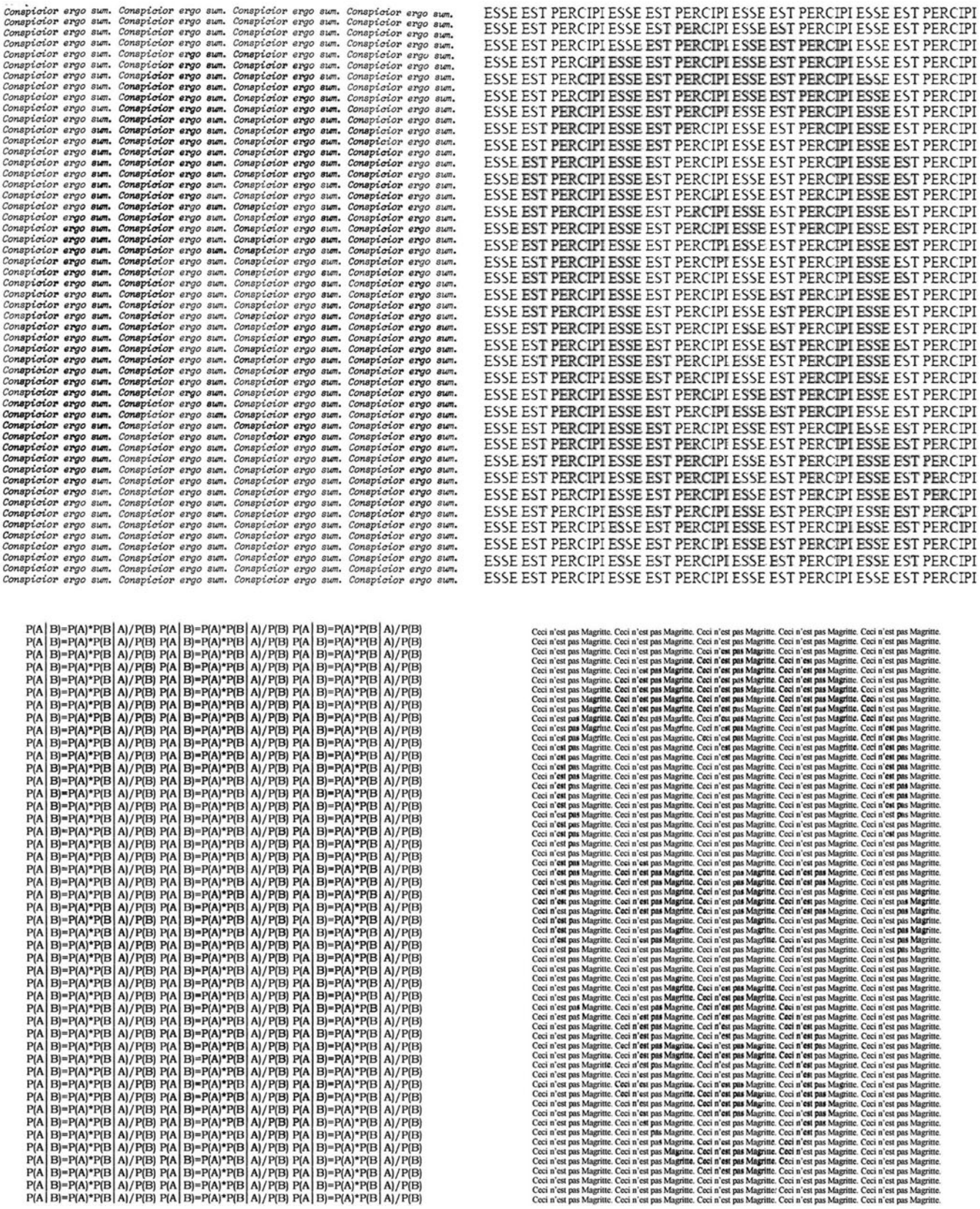


Figure X.113-3. Upper left: *Dual aspect*, the existence of René Descartes is determined not only by his rationality (*cogito*) but also by his visibility (*conspicior*). (From Wade, 1990.) Upper right: *Berkeley's esse* in which George Berkeley exists if he can be perceived. (From Wade, 2009.) Lower left: *Bayesian inference*, which is possibly a portrait of Rev. Thomas Bayes. (From Wade and Swanston, 2013.) Lower right: *Ceci est Magritte* is a portrait of the artist René Magritte hidden in text declaring that it is not him. (From Wade, 2011; see Wade, 2011, for a color version of this perceptual portrait.) (All illustrations by Nicholas Wade.)

primacy of thinking is matched in the illustration by that of perceiving. A similar theme is played with the portrait of Berkeley (1710), who wrote that *esse est percipi* ("to be is to be perceived"): he exists in the picture if he can be perceived. Formulae as well as text can carry portraits, and some of these are assuming increased prominence in perceptual research. Berkeley introduced notions of inference in visual perception, and they now find a new champion in an old mathematician, who is shown in Figure X.113-3. Bayes provided a formulation of inference that is widely used in perceptual theories. According to Bayesian theorists, there is always some prior probability of encountering certain events and the weighting of these is modified by perception. There is no known authenticated portrait of Bayes, but the one used in producing this illustration is most commonly presented as a likeness of him. The equation $P(A \mid B) = P(A)^* P(B \mid A)/P(B)$ is Bayes' theorem. In the context of perceptual theory, the terms are (A) hypothesis and (B) evidence regarding its validity. Thus the probability that the hypothesis will be supported by the evidence is determined by the prior probability P(A) and the ensuing perception. If the hypothesis is supported the prior is strengthened and vice versa. The fourth portrait in Figure X.113-3 is of Magritte, who played subtle games with words and images. Magritte appreciated that both are symbols of their referents, and he set them in opposition in his painting *The Perfidy of Images*: beneath a painted pipe was the text *Ceci n'est pas une pipe*. This distinction between words and images has been referred to as one between spatialized and stylized images (Wade, 1990). Spatialized images are what we commonly refer to as pictures, be they painted, printed, or flickering on a screen. Stylized images, as in written scripts, are obviously arbitrary and dependent on convention. They present us with indirect references to objects and ideas. They are stylized images that usually have little spatial similarity to the things they signify. Magritte set these two forms of symbolic representation in conflict with one another, and his discordant tune is played upon his own facial features in Figure X.113-3.

A similar combination of high and low spatial frequency image content for faces has been developed by Schyns and Oliva (1994, 1999); they refer to them as "hybrid faces." Unlike the concealed detectability of faces in Figures X.113-1, X.113-2, and X.113-3, the hybrid faces are readily seen as faces but their identities are defined by the high and low spatial frequencies they contain. As with the concealed portraits in Figures X.113-1, X.113-2, and X.113-3, the initially unseen hybrid face becomes visible when the sharply defined elements in the image are defocused or vice versa.

Gestalt principles of perceptual organization can operate to conceal as well as reveal. This applies to figure–ground and interpretative ambiguities (Wade, 2014). The alternative interpretation can be made difficult to recognize, and this has been an artistic enterprise for centuries (Mannoni et al., 2004). Elements of patterns can be grouped in many ways, according to the principles formulated by Wertheimer (1922, 1923). In general, the principles result in the emergence of a "whole" from the "parts" of a global organization from the local elements. They can also operate to conceal a global image in the local elements, as is the case in Figure X.113-4. This is a rare example of a truly digital image for two reasons. On the one hand, it is comprised solely of the typed digits 0 and 1. On the other hand, the digits depicted are digits—the metaphorical finger of God passing the spark of life to Adam's finger. The hands are derived from a detail of Michelangelo's fresco on the ceiling of the Sistine Chapel in the Vatican.

Other aspects of Gestalt organizing principles can be manipulated to conceal images, some of which are shown in Figure X.113-5. These include incomplete images, symmetry, and orientation. The general features of the pictorial images will be readily seen, but the recognition of their identities might prove more puzzling. In these cases it is

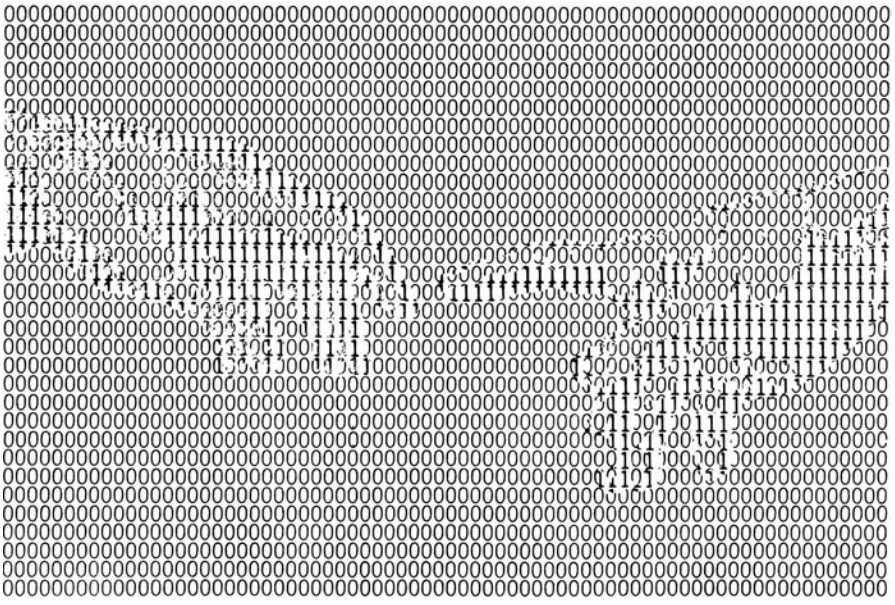


Figure X.113-4. *Digital picture* by Nicholas Wade. Detail of the hand of God passing the spark of life to Adam, depicted in digital form.

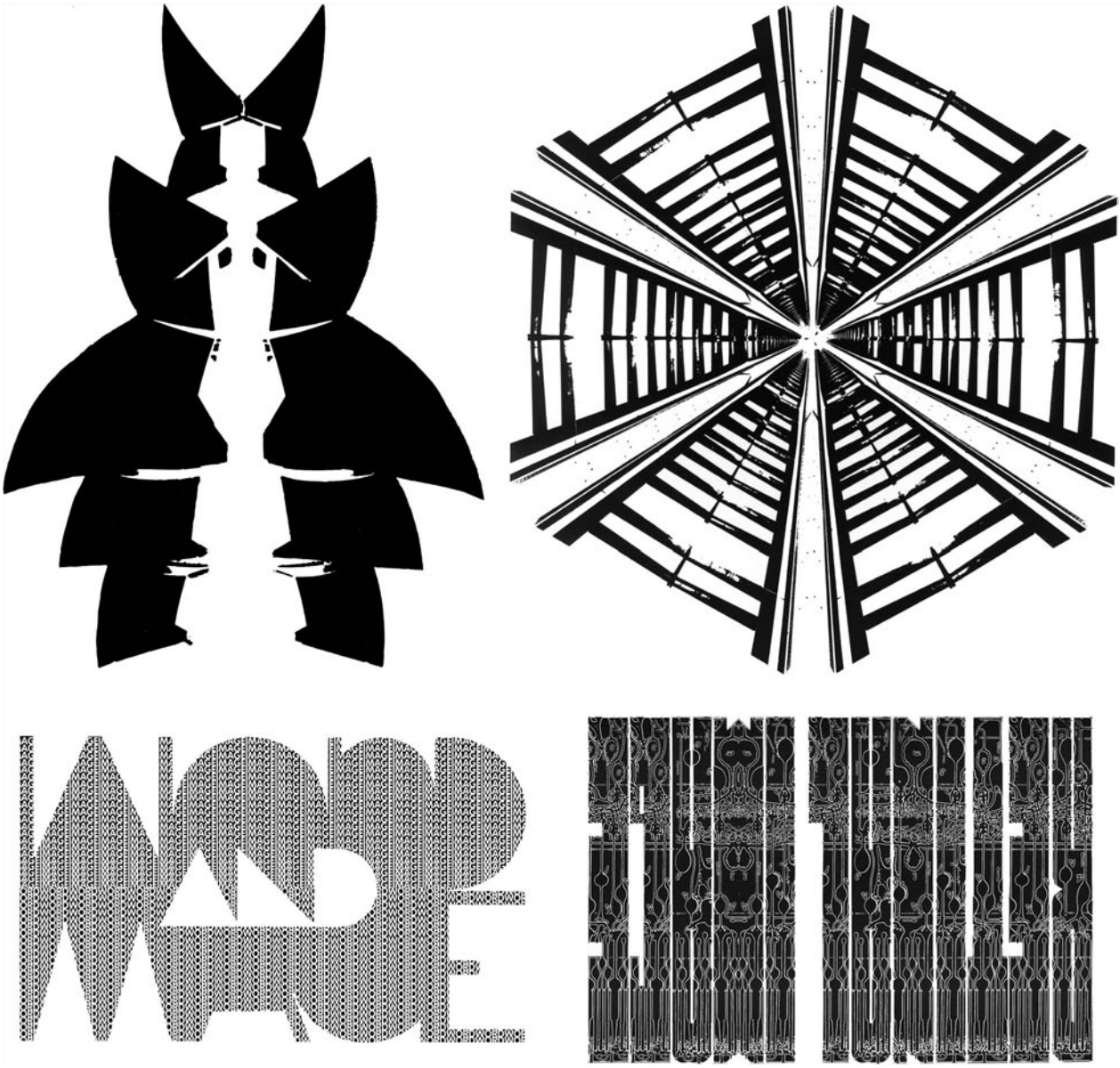


Figure X.113-5. Four figures that might be difficult to recognize at first, due to the operation of orientation and symmetry. Upper left: *Sydney Opera House* rendered in high contrast, symmetrically reflected and rotated 90°. Upper right: *A kaleidoscopic image of the Tay Road Bridge;* a photograph of the bridge was rendered in high contrast then rotated and reversed so that the 30° sectors are repeated 12 times. Lower left: *Word and image*; perception of the unconventional letter shapes defining these words is further disrupted by the clearly typed letters forming words that contradict those containing them. Lower, right: *Inverted retinal image*; the letters of the words retinal image are themselves inverted and reversed, containing within them a diagrammatic representation of retinal microscopic structure. (All illustrations by Nicholas Wade.)

not faces that are being concealed but either more concrete or more abstract aspects of the pictorial referents. A fundamental feature of Gestalt grouping that is often overlooked is orientation. The positioning of local and global features, with respect to the observer or to gravity, provides a signal source of segregation and recognition. Inversion adversely affects recognition not only of faces but also of text, as Rock (1973) remarked.

Artists have manipulated these dimensions of pictorial ambiguity for centuries, but many of the phenomena have been rediscovered in the confines of the visual laboratory (Wade, 2007). For example, Köhler's (1940) observation that pictures of inverted faces are difficult to identify was not a novelty for artists; examples can be found from Roman art onward (Wade, 2012). It has provided a method of hiding politically or socially unacceptable images, or in engaging in the license offered by pictorial representation. In the 17th century, an Italian artist named Giuseppe showed that it was even possible to cross the species and gender barriers by suitable graphical inversions (Wade & Nekes, 2005). Thompson (1980) made a subtle variation on this theme: inverting the eyes and mouth while maintaining the upright orientation of the head has a profound effect on the appearance of the face that is not matched by inverting the whole configuration. This manipulation has been widely mimicked (Wade, 2008).

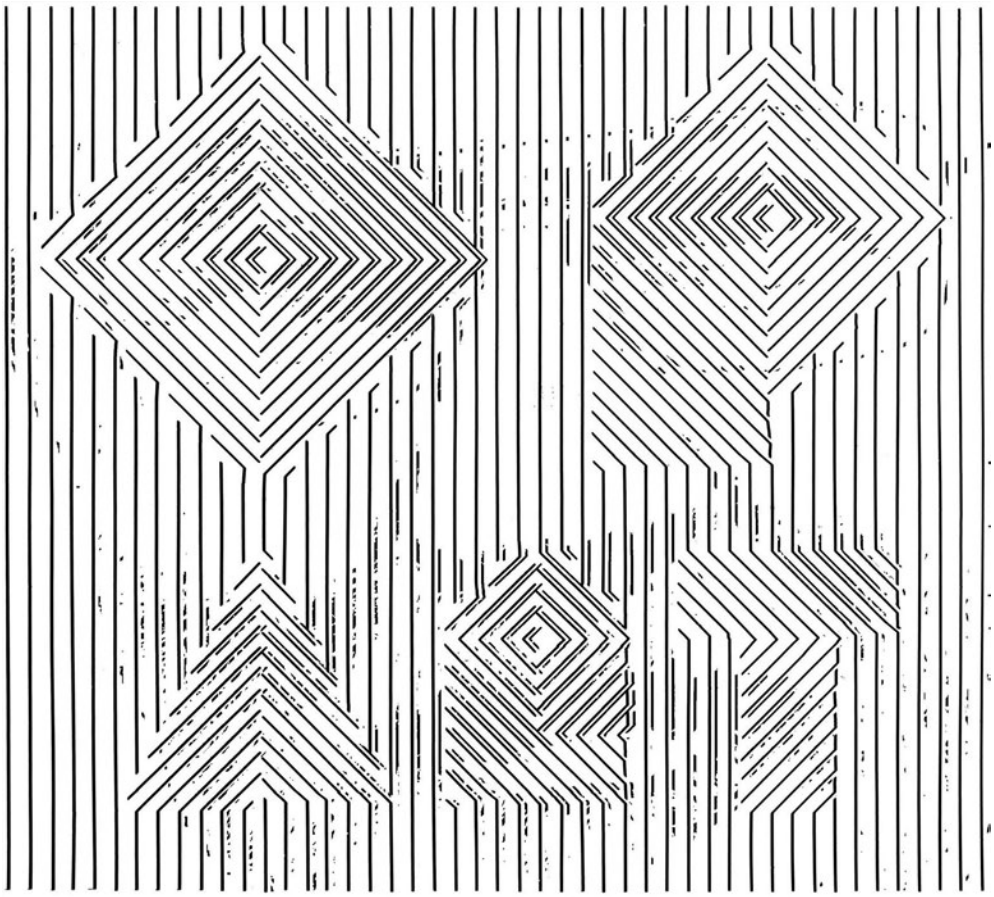


Figure X.113-6. *Op artistry* by Nicholas Wade.

There are many other methods for concealing images both in two and three dimensions. One of the driving factors in the development of random-dot stereograms was the desire to present disparities without object recognition, so that depth perception could be determined without object recognition (see Wade & Ono, 2012). Autostereograms serve a similar function because they appear initially flat and formless but they can create complex impressions of depth that articulate over time when suitably viewed.

Two puzzles are presented in the final illustration (Figure X.113-6). One concerns reading words within it and the other is a portrait that is carried by it. The words are defined by idiosyncratic letter shapes that speak to the enterprise that is embraced by them; the components of the face are readily detectable (unlike those in Figs. X.113-1, X.113-2, and X.113-3) but they are presented at a spatial scale that initially conceals their recognition.

REFERENCES

Baltrusaitis, J. (1976). *Anamorphic art* (W. J. Strachan, Trans.). London: Chadwyck-Healey.

Berkeley, G. (1710). *A treatise concerning the principles of human knowledge*. Dublin: Pepyat.

Bruce, V., & Young, A. (1998). *In the eye of the beholder: The science of face perception*. Oxford: Oxford University Press.

Calder, A. J., Rhodes, G., Johnson, M. H., & Haxby, J. V. (Eds.). (2011). *The Oxford handbook of face perception*. Oxford: Oxford University Press.

Descartes, R. (1637). *Discours de la méthode*. Leyden: Maire.

Gilad, S., Meng, M., & Sinha, P. (2009). Role of ordinal contrast relationships in face encoding. *Proceedings of the National Academy of Sciences USA*, *106*, 5353–5358.

Howard, I. P. (1982). *Human visual orientation*. Chichester, UK: Wiley.

Howard, I. P., & Templeton, W. B. (1966). *Human spatial orientation*. London: Wiley.

Köhler, W. (1940). *Dynamics in psychology*. London: Faber & Faber.

Leeman, F., Elffers, J., & Schuyt, M. (1976). *Hidden images: Games of perception, anamorphic art, illusion*. New York: Abrams.

Mannoni, L., Nekes, W., & Warner, M. (Eds.). (2004). *Eyes, lies and illusions*. London: Hayward Gallery.

Purkinje, J. (1825). *Beobachtungen und Versuche zur Physiologie der Sinne: Neue Beiträge zur Kenntniss des Sehens in subjectiver Hinsicht*. Berlin: Reimer.

Rock, I. (1973). *Orientation and form*. London: Academic Press.

Schyns, P., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science*, *5*, 195–200.

Schyns, P., & Oliva, A. (1999). Dr. Angry and Mr. Smile: When categorization flexibility modifies the perception of faces in rapid visual presentations. *Cognition*, *69*, 243–265.

Thompson, P. (1980). Margaret Thatcher: A new illusion. *Perception*, *9*, 483–484.

Wade, N. (1990). *Visual allusions: Pictures of perception*. London: Lawrence Erlbaum.

Wade, N. (1992). Faces of psychology. *Perception*, *21*(Suppl. 1), 1–107.

Wade, N. (1995). *Psychologists in word and image*. Cambridge, MA: MIT Press.

Wade, N. J. (2007). Artful visions. *Spatial Vision*, *21*, 27–53.

Wade, N. J. (2008). The Thatcherisation of faces. *Perception*, *37*, 807–810.
Wade, N. J. (2009). Berkeley's confused vision. *Perception*, *38*, 475–478.
Wade, N. J. (2010). Literal visions. *International Journal of Arts & Technology*, *3*, 251–274.
Wade, N. J. (2011). Eye contricks. *i-Perception*, 2, 486–501.
Wade, N. J. (2012). Artistic precursors of Gestalt principles. *Gestalt Theory*, *34*, 329–348.
Wade, N. J. (2014). Eye-conographs. *Art & Perception*, *2*(1–2), 3–10.
Wade, N. (2016). *Art and illusionists.* Heidelberg: Springer.
Wade, N. J., & Nekes, W. (2005). The two faces of Rex Whistler (1905–1944). *Perception*, *34*, 639–644.
Wade, N. J., & Ono, H. (2012). Early studies of binocular and stereoscopic vision. *Japanese Psychological Research*, *54*, 54–70.
Wade, N. J., & Swanston, M. (2013). *Visual perception: An introduction* (3rd ed.). London: Psychology Press.
Wertheimer, M. (1922). Untersuchungen zur Lehre von der Gestalt: I. Prinzipielle Bemerkungen. *Psychologische Forschung*, *1*, 47–58.
Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt: II. *Psychologische Forschung*, *4*, 301–350.

PART XI
Multisensory

Chapter 114
The Three-Dimensional Necker Cube

Nicola Bruno

The Necker cube is a widely known example of a reversible figure. Perceptual reversals were first observed in engravings of crystals by the Swiss geologist Louis Albert Necker (1832). Although Necker's engravings were not exactly of regular cubes, the figure is now typically drawn as in Figure X.114-1, which can be perceived in two alternative arrangements of a three-dimensional (3D) cube. Although both percepts are compatible with the proximal stimulus as provided by the drawing, that involving a cube resting on a horizontal support surface (Fig. X.114-1, left) is perceived more readily, presumably due to assumptions about typical viewpoints and support surfaces (Troje & McAdam, 2010). Fixations also tend to influence the perceived arrangement (Kawabata, Yamagami, & Noaki, 1978; Toppino, 2003). Under prolonged observation, most observers experience spontaneous reversals, although this may require some time if observers are not warned that the percept might change (up to 1 to 3 min according to one estimate; Girgus, Rock, & Egatz, 1977). The dynamics of reversals has been the object of numerous studies (e.g., Borsellino, De Marco, Allazetta, Rinesi, & Bartolini, 1972), which have considered it as a reflection of perceptual processing in the face of projective ambiguity in the proximal stimulus. This feature of the demonstration, perceptual changes despite unchanging optical stimulation, is presumably one of the drivers of the continuing fascination with the Necker cube. A Google search performed while revising this chapter returned about 140,000 hits, ranging from scientific discussions to popular science, simple pictures to more complex animations and applets.

Shortly after Necker's publication, Wheatstone (1838) described reversals that could occur with 3D skeleton objects (see Wade et al., 2010). However, the fact that reversals can and do occur with a 3D Necker cube when this is observed monocularly and with the head stationary has remained much less widely known. That this can happen is perhaps not surprising, given that these conditions of observation remove binocular and motion information, reducing the information in the proximal stimulus to what is normally available in a two-dimensional (2D) picture if care is taken to use black wires against a white background and to minimize luminance differences. The 3D version is, however, more interesting for several reasons. As we shall see shortly, the 3D Necker cube turns out to be a surprisingly rich model for phenomenological analysis and psychophysical experimentation, with implications for theoretical issues such as the definition of visual illusions, the role of global 3D interpretations in the integration of local sensory signals, and the exploratory and multisensory nature of perceptual processes.

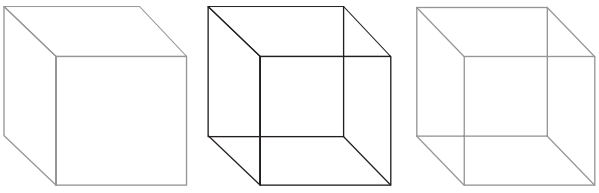


Figure X.114-1. Standard representation of the 2D Necker cube (black) and its two possible perceived depth arrangements (grey).

A first interesting feature of the three dimensional Necker cube lies in the nature of its 3D interpretations. Reversals of the 3D model imply a departure from veridical perception: when seen in the "true" depth arrangement, the cube appears indeed as a cube, but when seen in the "illusory" reversed arrangement, it appears as a truncated pyramid. This distinction between a veridical and an illusory perceived structure does not apply to the 2D case, as there is no external object there that can be taken as the "true" state of affairs. Indeed, it could be argued that the external object is, in this case, simply a set of connected lines on white paper. Taken this way, in the 2D demonstration both perceived 3D arrangements should be defined as illusions. Such illusions are interesting in themselves, as they testify to the strong tendency of human perception to interpret flat drawings as depicting a volume. This tendency is well worth further analysis. However, given that the focus of this chapter is on the 3D cube, I do not delve further on this issue.

Further observation of the 3D cube reveals a second interesting feature. Perception is active, and when we observe an object we typically move to explore multiple views. When doing this with the 3D cube, it is possible to notice another striking difference between the real and the illusory depth arrangement. When perceived in the real arrangement, the cube remains stationary (as it indeed is). When perceived in the illusory arrangement, conversely, the cube appears to rotate as if it followed the viewer. The illusory movement has been attributed

to the operation of perceptual mechanisms that stabilize the perceived world during active vision. As we move our body, the projection of the environment on our retinas is subjected to various changes that depend on spatial relationship, fixations, and type of movement. Yet we perceive a stable world: A sophisticated brain mechanism "knows" how we have moved our eyes and our body and somehow manages to discount the optical changes that can be attributed to these movements.

This mechanism correctly stabilizes the cube when seen in the true arrangement, as the retinal motions can be interpreted as due to changes in 3D positions that would be expected given the movement of the observer (Fig. X.114-2, left column). When the cube is seen in the reversed illusory arrangement, however, the interpreted changes in relative 3D position cannot be due to the movement of the observer (right column), and this causes an attribution of the retinal motions to movement of the object. The motion is illusory, but the illusion arises from a mechanism that would produce a veridical percept if the depth arrangement had been perceived veridically. Thus the illusion may be taken as evidence that the process that stabilizes the perceived world during active vision is not a simple comparison between the internal representation of the movement of the observer and the corresponding retinal change but rather involves "intelligent" computations that take into account the perceived 3D structure in a highly sophisticated manner (Rock, 1982).

But perhaps the most surprising effect to be observed using a 3D Necker cube is what happens when one holds a small-scale model of the cube in one's hands. Under these conditions, one might expect that the information about object structure obtained by touching would prevent any reversal to occur. This expectation is quite reasonable, given that the sensory signals from touch receptors in the hand are always consistent with the true interpretation (a cube) and with the veridical depth order of the cube faces. As anyone can readily confirm by constructing a suitable wireframe cube, however, the expectation turns out to be wrong. If we take care to keep our hands on the outer parts of the cube, avoiding placing our fingers inside the wireframe (e.g., by holding the cube as in Fig. X.114-3, middle left panel), and we have the patience to wait a few seconds, reversals of the cube will eventually take place. The resulting experience is striking: at times, the object will seem to undergo nonrigid deformations; at other times, we will have the impression that our wrists are bent at impossible angles, consistent with the visually reversed shape instead of the haptically felt one. To the best of my knowledge, this surprising observation was first reported by Shopland and Gregory (1964), who used a self-luminous wireframe cube observed in the dark to minimize vision of the hands. However, it turns out that this is not strictly necessary, as reversals occur even if the hands are fully visible (Bruno, Bertamini, Jacomuzzi, & Meyer, 2007).

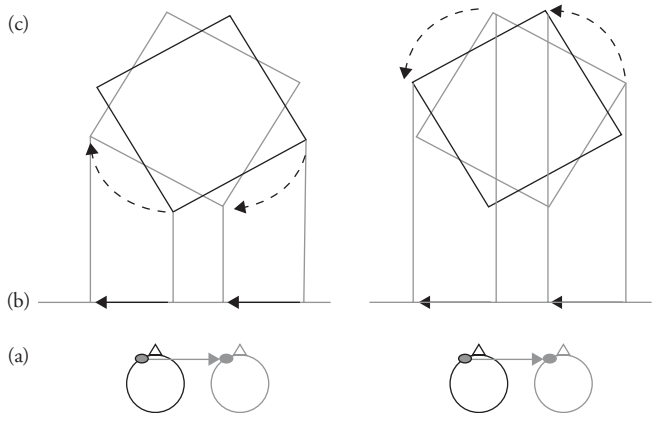


Figure X.114-2. The movement of a monocular observer (a), the retinal displacements for the cube contours (b), and the corresponding interpretations as 3D changes of position relative to the observer (c) when the cube is perceived in its true depth arrangement (left) or in the illusory, reversed arrangement (right). (Credits should be to Nicola Bruno.)

At first blush, reversals with a handheld cube seem to suggest visual dominance over touch (visual "capture"), but more careful measurements indicate that they instead reflect a multisensory process that takes into account both the visual and the tactile signals. Evidence for this conclusion comes from studies of reversal frequency and of the duration of the perceived alternatives. For instance, Shopland and Gregory (1964) compared reversal frequencies when seeing and touching the cube and when simply seeing it. They found that the rate of reversals was significantly reduced in the former condition in comparison to the latter. Ando and Ashida (2003) compared the average durations of multisensory consistent percepts (which occur when the cube is seen in its true depth arrangement) with that of inconsistent ones (when the cube is seen in its illusory arrangement). They found that the average duration of consistent percepts was significantly longer than that of the inconsistent ones. Bruno et al. (2007) compared average durations of the veridical and the illusory percept in three conditions (see Fig. X.114-3a): holding the cube with a pincer grip (baseline: no shape information), holding it with the hand palms on two opposite vertices (hand stationary: passive touch), and holding it while systematically moving one of the hands around one of the vertices (hand moving: active touch). In all three conditions they observed longer durations of the veridical percepts, but in the baseline condition this difference was small and statistically not significant; conversely, the difference was larger and statistically significant in the hand stationary condition, and even larger and again significant in the hand moving condition. These findings suggest that although reversals occur despite conflicting information from the hands, this does not reflect complete visual dominance. Instead, the information from touching the cube interacts with the visual information, making the reversals less likely in part also as a function of the nature of the sensory signal.

This pattern of effects makes the handheld Necker cube an interesting model to investigate multisensory processes during extended periods of exploration. Perception is active and multisensory, and we often look at an object while we feel it with our hand. Perceptual exploration by eye and hand extends in time, and as exploration progresses new information may require changing how the sensory signals are treated. For instance, signals that were in agreement and were merged may suddenly disagree. One would expect that multisensory merging would take this disagreement into account somehow, but collecting data that might provide information on such adaptive processes is not easy. One way to collect such data is to exploit spontaneous reversals in a configuration like the 3D cube. In ordinary vision, of course, objects do not usually reverse in

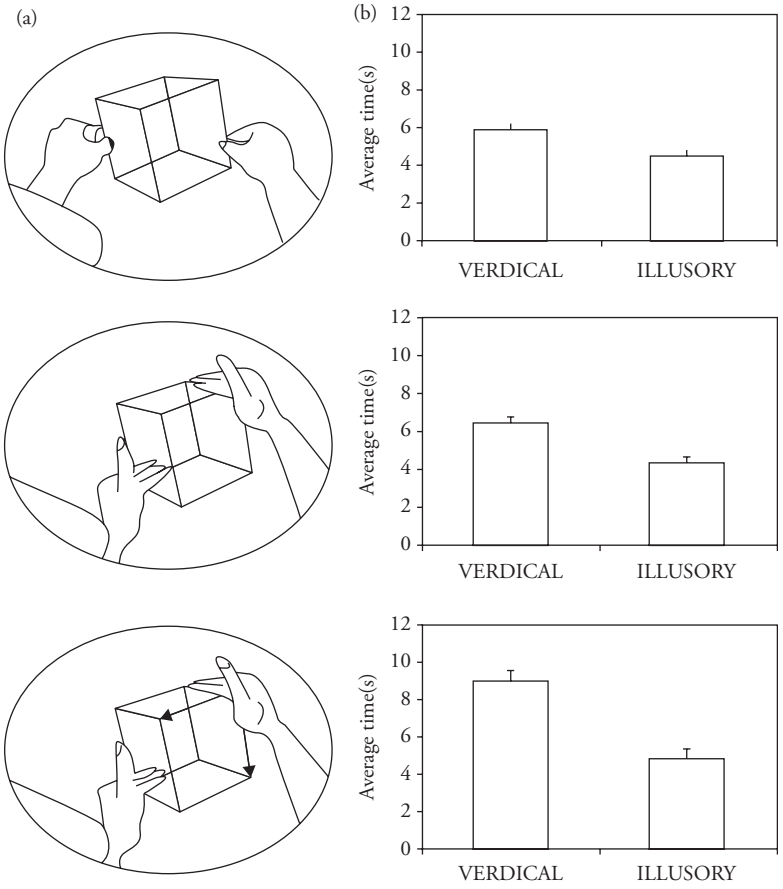


Figure X.114-3. (a) The baseline, hand stationary, and hand moving conditions of Bruno et al.'s (2007) study. (b) The corresponding durations of the true and illusory percepts, averaged across 17 participants and plotted with one structural equation modeling error bars. (Adapted from Bruno et al., 2007.)

depth. Still, the dynamics of reversals provide interesting information about multisensory interactions when vision and touch agree or when they begin to conflict.

Using a novel paradigm that they developed specifically to study reversal dynamics, Bruno et al. (2007) asked participants to continuously look at the handheld cube while they alternated periods with a stationary hand and periods when they moved the hand. During several 1-min sessions, participants heard signals that instructed them to start moving or to stop at random times and continuously reported reversals. By studying not only how often reversals occurred but also *when* they occurred, Bruno et al. were able to compute cumulative probabilities of either type of reversal (from the veridical to the illusory percept or vice versa) given either hand state (stationary or moving) as a function of the time elapsed since the previous change in hand state. This resulted in four curves that, not surprisingly, indicated that the likelihood of a reversal was relatively higher soon after the change in hand state and tended to taper off after about 7 to 8 s when, by design, another hand state was increasingly likely to occur. Closer inspection of the curves, however, revealed a striking difference between the curve corresponding to the veridical-to-illusory reversals *after the hand began to move* and the other three curves. Whereas the other three curves showed a continuous, monotonic increase in the cumulative probability, veridical-to-illusory reversals after the stationary-to-moving transition increased normally up to about 1 to 2 s, stopped increasing for approximately another 2 s, and only then began again to increase monotonically (Fig. X.114-4). Fitting the curves with cumulative gamma functions (Borsellino et al, 1972; Mamassian

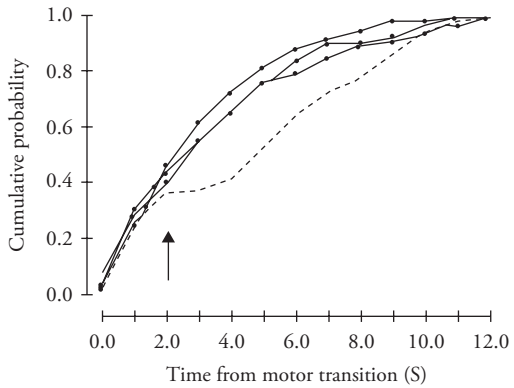


Figure X.114-4. Cumulative probabilities of a given reversal (veridical to illusory or illusory to veridical) after a given change in hand state (stationary to moving and moving to stationary). Dashes identify the curve for the veridical-to-illusory reversals after a stationary-to-moving hand change. Note that after about 2 seconds and for about 2additional seconds, it was effectively impossible for participants to experience this kind of reversal. (Adapted from Bruno et al., 2007.)

& Goutcher, 2005) confirmed the difference between the curves, and a similar pattern was later observed in follow-up experiments (Bertamini, Masala, Meyer, & Bruno, 2010).

Bruno et al. (2007) interpreted this pattern as evidence for a "vetoing" window in the multisensory interaction between vision and haptics during the continuous perceptual exploration of the object. As the hand began to move, it acquired more accurate information about the true shape of the object. This made it effectively impossible for participants to flip the cube to its illusory interpretation, but only after some time and within a specific temporal window. Interestingly, this window was fairly narrow, effectively preventing illusory reversals only for about 1 to 2 s. This may suggest that, within extended perceptual explorations, the system is not continuously measuring the quality of sensory signals but is instead built to be sensitive to changes in this quality. When such transient changes occur within a sensory channel, this channel exerts a strong influence on the multisensory percept (Shimojo & Shams, 2001), but after this change has been registered adaptation takes place and eliminates the vetoing effect. Vetoing effects are a form of multisensory interaction that is still relatively unexplored, and they may prove a fruitful area of investigation in the future.

REFERENCES

Ando, H., & Ashida, H. (2003). Touch can influence visual depth reversal of the Necker cube. *Perception*, *32*, 97.

Bertamini, M., Masala, L., Meyer, G., & Bruno, N. (2010). Vision, haptics, and attention: New data from a multisensory Necker cube. *Perception*, *39*(2), 195–207.

Borsellino, A., De Marco, A., Allazetta, A., Rinesi, S., & Bartolini, B. (1972). Reversal time distribution in the perception of ambiguous stimuli. *Kybernetik*, *10*, 139–144.

Bruno, N., Jacomuzzi, A. C., Bertamini, M., & Meyer, G. (2007). A visual-haptic Necker cube reveals temporal constraints on intersensory merging during perceptual exploration. *Neuropsychologia*, *45*, 469–475.

Kawabata, N., Yamagami, K., & Noaki, M. (1978). Visual fixation points and depth perception. *Vision Research*, *18*, 853–854.

Mamassian, P., & Goutcher, R. (2005). Temporal dynamics in bistable perception. *Journal of Vision*, *5*, 361–375, http://journalofvision.org/5/4/7/

Necker, L. A. (1832). Observations on some remarkable phenomena seen in Switzerland: And on an optical phenomenon which occurs on viewing a crystal or geometrical solid. *The London and Edinburgh Philosophical Magazine and Journal of Science*, *1*, 329–337.

Rock, I. (1982). *The logic of perception*. Cambridge, MA: MIT Press.

Shimojo, S., & Shams, L. (2001). Sensory modalities are not separate modalities: Plasticity and interactions. *Current Opinions in Neurobiology*, *11*(4), 505–509.

Shopland, J. C., & Gregory, R. L. (1964). The effect of touch on a visually three-dimensional figure. *The Quarterly Journal of Experimental Psychology*, *16*, 66–70.

Toppino, T. (2003). Reversible-figure perception: Mechanisms of intentional control. *Perception & Psychophysics*, *65*, 1285–1295.

Troje, N. F., & McAdam, M. (2010). The viewing-from-above bias and the silhouette illusion. *i-Perception*, *1*, 143–148.

Wheatstone, C. (1838). Contributions to the physiology of vision: Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision, *Philosophical Transactions of the Royal Society of London*, *128*, 371–394.

Chapter 115
The McGurk Effect and the Primacy of Multisensory Perception

James W. Dias, Theresa C. Cook, and Lawrence D. Rosenblum

INTRODUCTION

The McGurk effect occurs when an audible speech syllable is presented in synchrony with a visible face articulating a different speech syllable. The resulting perception is of a single "heard" syllable that has been visually influenced in some way (McGurk & MacDonald, 1976). In a classic example, perceivers presented with an audible "ba" while seeing a face articulating "ga" will report *hearing* a clear, unitary "da" syllable. The McGurk effect reveals that multisensory integration can be thorough, impervious to outside influences, and primary, beginning at the very earliest levels of processing. For these reasons, the McGurk effect has come to be regarded as a prototypical demonstration of multimodal perception. When investigating cross-modal integration in any domain, researchers often ask: Is this integration as complete as the McGurk effect?

Because of its robust nature, the McGurk effect has also been used as a tool to address issues outside the domain of multisensory speech integration. These issues include the degree to which hearing listeners incorporate lip-read information (Baart & Vroomen, 2010; Rosenblum & Saldana, 1992; 1996; Andersen, Tiippana, Laarni, Kojo, & Sams, 2009; Vatikiotis-Bateson, Eigsti, Yano, & Munhall, 1998); whether speech is perceived in a specialized manner (e.g., Saldana & Rosenblum, 1993); the relation between face and visual speech perception (Green, Kuhl, Meltzoff, & Stevens, 1991; Rosenblum, Yakel, & Green, 2000; Sekiyama, 1998; Thomas & Jordan, 2002; Walker, Bruce, & O'Malley, 1995); and the perceptual primitives of general speech perception (e.g., Fowler, 1986; Green, 1998; Rosenblum, 2005, 2008).

More recently, McGurk-type cross-modal speech influences have been found to extend beyond the visual domain to include influences of tactile (Fowler & Dekle, 1991; Gick & Derrick, 2009) and kinesthetic (Ito, Tiede, & Ostry, 2009; Sams, Mottonen, & Sihvonen, 2005) speech information on heard speech. These and other recent findings have motivated a conception of the speech mechanism that is supremely multisensory and able to extract all linguistically relevant articulatory information, with little regard for sensory modality. In turn, this understanding of the speech mechanism as designed around event rather than modality-specific primitives has influenced recent conceptions of the perceptual brain in general. In this sense, the McGurk effect has evolved from a simple, albeit compelling, audiovisual speech illusion to an effect that may have critical implications for the nature of the brain.

In this chapter, we (a) review research on the McGurk effect, showing its strength, automaticity, and flexibility with regard to sensory modality; (b) discuss how the effect has revealed a speech mechanism that treats all relevant crossmodal information similarly; and (c) examine how this research has helped motivate a reconceptualization of the perceptual brain for which multisensory information is central. Throughout, explanations of the McGurk effect are addressed in the context of the empirical data. We begin by describing the original discovery of the effect. (For a more recent review of the relevant literature, see Rosenblum, Dias, & Dorsi, 2016.)

THE STRENGTH, AUTOMATICITY, AND FLEXIBILITY OF THE MCGURK EFFECT

McGurk and MacDonald (1976) first demonstrated the audiovisual speech illusion by dubbing auditory syllables onto videos of a face articulating syllables with different leading consonants. When asked to watch these videos and state what they *heard*, perceivers reported coherent syllables whose leading consonants were strongly influenced by the visual information. Depending on which syllables were involved, perceivers either heard a fusion of the consonants (e.g., auditory "ba-ba" + visual "ga-ga" = heard "da-da"), a combination of the consonants (auditory "ga-ga" + visual "ba-ba" = heard "bda-bda" or "bga-bga"), or (in subsequent studies) a visually dominated consonant (auditory "ba" + visual "va" = heard "va"). (Other combinations of consonants show no visual influence.)

McGurk and MacDonald (1976) postulated that the nature of the visual influence depends on the ways that the auditory and visual signals share common articulatory information. For example, auditory "ba" + visual "ga" = sounded like "da" because the visual signal provides articulatory information for both "ga" and "da" and the auditory signal provides articulatory information for "ba" and "da." Thus the *fused* "da" audiovisual percept results from the common segment information for "d" in the audio and visual signals (MacDonald & McGurk, 1978; McGurk & MacDonald, 1976).

Since its original demonstration, substantial research has revealed the breadth, durability, and automaticity

of what has become known as the McGurk effect. The effect has been shown to work with many different consonant combinations (e.g., MacDonald & McGurk, 1978; McGurk & MacDonald, 1976), as well as, to a more subtle degree, vowel combinations (e.g., Massaro & Cohen, 1993; Summerfield & McGrath, 1984). With regard to its durability, the McGurk effect remains effective when the auditory and visual components are spatially separate (Bertelson, Vroomen, Wiegeraad, & de Gelder, 1994), when observers are instructed to direct their gaze to the side of the face (Richler, Hove, & Munhall, 2003), and when the visual and auditory components are temporally offset to a noticeable degree (Munhall, Gribble, Sacco, & Ward, 1996; Soto-Faraco & Alsius, 2009). The effect seems cognitively impenetrable to a large extent: perceivers will still report a McGurk effect even when they know the auditory and visual components of the audiovisual stimulus are incongruent (e.g., Bertelson & De Gelder, 2004; Colin et al., 2002; Soto-Faraco & Alsius, 2009) and are provided by talkers of different genders (Green et al., 1991).

Further demonstrations of the effect's strength include its effectiveness when the auditory (Andersen, Tiippana, Laarni, Kojo, & Sams, 2009) or visual (Thomas & Jordan, 2002) components of the stimulus are degraded. For example, perceivers experience the McGurk effect even when facial features are removed from the video, leaving only 13 small white points that move with the speaker's articulation (Rosenblum & Saldana, 1996). In this example, some subjects were never aware that they were seeing a face yet still showed a significant visible speech influence in their syllable identification. Based on this observation, it seems the McGurk effect works even when perceivers does not know that they are looking at a face.

The McGurk effect has been shown to work on perceivers of different native-language backgrounds. For example, the effect has been shown to work with perceivers who are native speakers of Japanese (Massaro, Cohen, Gesi, Heredia, Tzuzaki, 1993; Sekiyama & Tohkura, 1991), Finnish (Sams et al., 1998), German, and Spanish (Fuster-Duran, 1996), although the segments that are most effective depend on the consonant inventory of the perceiver's language. The effect also works with prelinguistic infants as young as four months of age, using gaze habituation and habituation-test techniques for both visually dominated and fused crossmodal syllables (Burnham & Dodd, 2004; Desjardins & Werker, 2004; Rosenblum, Schmuckler, & Johnson, 1997).

These results on infants bear on the question of whether associative experience between auditory and visual signals is needed for the effect to occur (see Green, 1998). From some accounts, the McGurk effect, and audiovisual speech integration in general, is a consequence of having experience associating visible speech input with auditory speech representations (e.g., Massaro, 1998). The finding that four-month-old infants display a McGurk effect suggests that relatively little such experience is needed. Along these lines, there is evidence that newborn infants seem sensitive to phonetic correspondences across auditory and visual speech streams (e.g., Aldridge, Braga, Walton, & Bower, 1999; Aronson & Rosenbloom, 1971; Walton & Bower, 1993). While not evidence for crossmodal *integration,* as such, these neonate results suggest that associative experience is not necessary for awareness that the same speech gesture can be conveyed through auditory and visual modalities.

Other evidence that associative experience may not be necessary comes from adult studies showing integration from a modality that has not been previously used for speech perception. For example, a McGurk-type effect occurs when auditory information is paired with incongruent manual *touch* information, instead of visual information (Fowler & Dekle, 1991; Gick & Derrick, 2009). Fowler and Dekle found that touching a face as it articulates a speech syllable while simultaneously hearing an incongruent syllable induces a perceived fused syllable. More recently, Gick and Derrick found that silent puffs of air presented to the neck or hand presented synchronously with an auditory syllable "ba" can make the syllable sound like a "pa." The puffs of air, it is thought, act to provide information for the aspirated air that leaves the vocal tract during production of a voiceless "p," in contrast to a voiced "b."

In another example of nonvisual multisensory speech integration, manipulated somatosensory information from subjects' own speech articulators can influence heard speech. Ito and colleagues (2009) attached wires to the skin surrounding participants' mouths so that by electronically pulling on these wires, participants' mouths could be subtly positioned into different configurations. When these pulls were imparted synchronously with auditorily presented syllables, subjects' induced mouth positions had systematic influences on what they reported hearing. Similar findings have been observed when subjects were asked to silently mouth syllables incongruent with those they simultaneously heard (Sams et al., 2005).

These nonvisual examples of multisensory speech integration are significant for many reasons. First, they show that cross-sensory integration of speech is not limited to auditory and visual modalities. The speech mechanism seems able to integrate input from any modality that provides speech information. Second, in showing that integration can occur using a modality through which speech is not typically conveyed, the findings add to the developmental literature suggesting that little, if any, associative experience is needed for speech to integrate across modalities. It is unlikely that adults have experience perceiving speech from touching faces, or feeling puffs of air on the back of the neck, or having their mouths configured in a particular position. Still, all of these sources of information seem to readily integrate with auditorily presented speech.

Thus it is likely that the McGurk effect is not a result of simple associative experience between information across the modalities. This is not to say that experience itself does not influence the effect. Developmental research shows that there are changes in the strength of the effect throughout childhood (Dupont, Aubin, & Menard, 2005; McGurk & MacDonald, 1976), with a notable decrease when children are learning to read (Jerger, Damian, Spence, Tye-Murray, & Abdi, 2009). Also, as stated, while perceivers from all native-language backgrounds show some form of the effect, the particular segments that are most effective differ between languages (e.g., Sekiyama & Tohkura, 1993; Massaro, Cohen, Gesi, Heredia, Tzuzaki, 1993). Language experience is also relevant to observed influences of lexical and semantic context on the strength of the effect

(Brancazio, 2004). Still, while a perceiver's experience can have an influence on the strength of the McGurk effect, the basis of the integration itself is likely not dependent on associative experience, as the previously mentioned developmental, tactile, and somatosensory/kinesthetic demonstrations suggest. Instead, the effect may reflect a speech brain that naturally takes in all articulator-relevant information with relatively little concern for modality. Evidence for this explanation of the McGurk effect is discussed next.

THE MCGURK EFFECT DEMONSTRATES THAT THE SPEECH MECHANISM IS MULTISENSORY

The robust and automatic nature of multisensory integration in the McGurk effect has helped motivate a view of speech perception in which the perceptual primitives are not tied to any single modality. Instead, the speech function is thought to work with *articulatory* primitives—comprised of dynamic actions of the vocal tract—many of which can be extracted from multiple modalities (e.g., Fowler, 1986, 2004; Liberman & Mattingly, 1985; Rosenblum, 2008). In this sense the speech function is *amodal* in its operations as it searches for any information that is relevant for recovering a unified articulatory gesture.

Usually, crossmodal information will specify the same speech gesture. However, for McGurk effect stimuli, the speech mechanism is faced with trying to extract whatever common information is available in the incongruent articulations presented across the modalities. For example, an auditory "ba" + visual "ga" could provide a range of crossmodal information that ends up specifying a "da" articulation. This explanation of the effect's phonetic outcomes is reminiscent of that originally proposed by McGurk and MacDonald (1976; see previous discussion), with the important difference that the crossmodal influence occurs at the information *extraction* stage itself, before speech segments—or even speech features—have been established by the system.

In fact, there is both behavioral and neurophysiological evidence that crossmodal speech integration occurs exceedingly early and perhaps before speech features have been determined (e.g., Calvert, Campbell, & Brammer, 2000; Campbell, 2008; Green, 1998;; Rosenblum, 2005; Summerfield, 1987). For example, using a McGurk effect methodology, researchers have found that the aforementioned vocal air-burst feature that helps distinguish "b" and "p" can, itself, be perceptually shifted with visual information for "b" spoken at different rates (Green & Miller, 1985). With regard to neurophysiological research, multiple studies have revealed that visible speech can directly modulate auditory cortex (e.g., Calvert et al., 1997; Campbell, 2008; Pekkola et al., 2005) and even the earlier physiological stage of auditory brain stem (Musacchia et al., 2005). In addition, tests using McGurk-like stimuli have shown that changes in the visual component can actually override the auditory component in modulating auditory cortex (Callan, Callan, Kroos, & Vatikiotis-Bateson, 2001; Colin et al., 2002; Mottonen, Krause, Tiipana, & Sams, 2002; Sams et al., 1991). For example, changing the auditory component of an auditory "ba" + visual "va" (most always perceived as "va") to a "va" will *not* change auditory cortex activity. However, changing the visual component of auditory "ba" + visual "va" to "ba" (inducing a perceived "ba") *will* change auditory cortex activity. Thus auditory cortex reactivity to these McGurk effect stimuli follows the visual but not auditory component and is consistent with perceivers' reports of *hearing* visually influenced speech segments.

In sum, much of the behavioral and neurophysiological research on the McGurk effect suggests that crossmodal influences occur at an exceptionally early stage (however, see Bernstein, 2005; Jiang & Bernstein, 2011, for an alternative account). The neurophysiological results are also consistent with a speech function that acts to recover information about articulation with little regard for sensory modality. The aforementioned tactile and somatosensory/kinesthetic McGurk-type effects also support this notion. It is as if the speech mechanism searches for any information relevant to articulation and treats it similarly, whether extracted from auditory, visual, tactile, or kinesthetic sources.

THE MCGURK EFFECT HAS HELPED MOTIVATE A NEW PERSPECTIVE ON THE PERCEPTUAL BRAIN

The aforementioned finding that visual speech can modulate auditory cortex is important not only in its implications for multisensory speech perception but for perceptual psychology overall. This finding (Calvert et al., 1997) was the first (known to us) to show visual modulation of what was typically considered a sound-specific area in the human brain. Since then, crossmodal influences on cortical areas have been observed for nonspeech as well (e.g., Calvert et al., 2000; Macaluso, 2006; Shams & Seitz, 2008). These effects include visible nonspeech event modulation of auditory cortex (e.g., Bulkin & Groh, 2006; Kayser, Petkov, & Logothetis, 2008; Stekelenburg & Vroomen, 2007), tactile modulation of visual cortex (e.g., Amedi, Malach, Hendler, Peled, & Zohary, 2001; Merabet et al., 2008), and auditory modulation of visual cortex (e.g., Shams, Iwaki, Chawla, & Bhattacharya, 2005; Shams, Kamitani, Thompson, & Shimojo, 2001; Watkins, Shams, Tanaka, Haynes, & Rees, 2006).

The McGurk effect has also helped motivate behavioral research on sensory integration. Information from the visual modality can bias nonspeech auditory (e.g., Saldana & Rosenblum, 1993; Schutz & Lipscomb), proprioceptive (e.g., Harris, 1965; Hay, Pick, & Ikeda, 1965; Holmes & Spence, 2005; Lateiner & Sainburg, 2003; Lee & Aronson, 1974; Lee & Lishman, 1977), vestibular (e.g., Clark & Greybiel, 1966; Mergner & Rosemeier, 1998), and tactile (e.g., Botvinick & Cohen, 1998; Dopjans, Wallraven, & Bulthoff, 2009; Kinney & Luria, 1970; Konkle, Wang, Hayward, & Moore, 2009; Norman, Clayton, Norman, & Crabtree, 2008; Slater, Perez-Marcos, Ehrsson, & Sanchez-Vives, 2008) perception. Auditory information can bias visual judgments as well (e.g., Bertelson & Radeau, 1981; Sekuler, Sekuler, & Lau, 1997; Shams, Kamitani, & Shimojo, 2002; Walker & Scott, 1981).

Taken together, this research supports a conception of a general perceptual brain that is designed around multisensory input (e.g., Shimojo & Shams, 2001). Further, research showing neonate sensitivity to cross-sensory correspondences (see Walker-Andrews, 1997) suggests that, like

for speech, simple associative experience is not necessary for the perceptual brain to detect cross-sensory information. Instead, as for speech perception, general perceptual mechanisms may be much less modality differentiated than has been assumed. The perceptual brain seems more organized around event and object dimensions than sensory modality as such.

In sum, since its discovery, the McGurk effect has proven to be an impressively durable, automatic, and far-reaching intersensory illusion. Recent findings have demonstrated the effect's flexibility, showing that it extends beyond a simple audiovisual illusion such that tactile and somatosensory/kinesthetic information can also influence heard speech. It is likely all of these reasons have contributed to the effect being so central in motivating research and theorizing about a new view of multisensory perception and the perceptual brain.

ACKNOWLEDGMENTS

Work on this chapter was supported by National Institute on Deafness and Other Communication NIDCD Grant 1R01DC008957-01.

REFERENCES

Aldridge, M. A., Braga, E. S., Walton, G. E., & Bower, T. G. (1999). The intermodal representation of speech in newborns. *Developmental Science*, *2*, 42–46.

Amedi, A., Malach, R., Hendler, T., Peled, S., & Zohary, E. (2001). Visuo-haptic object-related activation in the ventral visual pathway. *Nature Neuroscience*, *4*(3), 324–330.

Andersen, T. S., Tiippana, K., Laarni, J., Kojo, I., & Sams, M. (2009). The role of visual spatial attention in audiovisual speech perception. *Speech Communication*, *51*, 184–193.

Aronson, E., & Rosenbloom, S. (1971). Space perception in early infancy: Perception within a common auditory-visual space. *Science*, *172*, 1161–1163.

Baart, M., & Vroomen, J. (2010). Do you see what you are hearing? Cross-modal effects of speech sounds on lipreading. *Neuroscience Letters*, *471*, 100–103.

Bernstein, L. E. (2005). Phonetic processing by the speech perceiving brain. In D. B. Pisoni & R. E. Remez (Eds.), *The handbook of speech perception* (pp. 79–98). Oxford: Blackwell.

Bertelson, P., & De Gelder, B. (2004). The psychology of multimodal perception. In C. Spence & J. Driver (Eds.), *Crossmodal space and crossmodal attention* (pp. 141–179). Oxford: Oxford University Press.

Bertelson, P., & Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Perception & Psychophysics*, *29*(6), 578–584.

Bertelson, P., Vroomen, J., Wiegeraad, G., & de Gelder, B. (1994). Exploring the relation between McGurk interference and ventriloquism. In *International Conference on Spoken Language Processing* (Vol. 2, pp. 559–562). Yokohama: Acoustical Society of Japan.

Botvinick, M., & Cohen, J. (1998). Ruber hands "feel" touch that eyes see. *Nature*, *391*, 756.

Brancazio, L. (2004). Lexical influences in audiovisual speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(3), 445–463.

Bulkin, D. A., & Groh, J. M. (2006). Seeing sounds: Visual and auditory interactions in the brain. *Current Opinion in Neurobiology*, *16*, 415–419.

Burnham, D. K., & Dodd, B. (2004). Auditory-visual speech integration by prelinguistic infants: Perception of an emergent consonant in the McGurk effect. *Developmental Psychobiology*, *45*(4), 204–220.

Callan, D. E., Callan, A. M., Kroos, C., & Vatikiotis-Bateson, E. (2001). Multimodal contribution to speech perception revealed by independent component analysis: A single-sweep EEG case study. *Cognitive Brain Research*, *10*, 349–353.

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C. R., McGuire, P. K., . . . David, A. S. (1997). Activation of auditory cortex during silent lipreading. *Science*, *276*, 593–596.

Calvert, G. A., Campbell, R., & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current Biology*, *10*(11), 649–657.

Campbell, C. (2008). The processing of audio-visual speech: Empirical and neural bases. *Philosophical Transactions of the Royal Society B*, *363*, 1001–1010.

Clark, B., & Greybiel, A. (1966). Factors contributing to the delay in the perception of the oculogravic illusion. *The American Journal of Psychology*, *79*(3), 377–388.

Colin, C., Radeau, M., Soquet, A., Demolin, D., Colin, F., & Deltenre, P. (2002). Mismatch negativity evoked by the McGurk-MacDonald effect: A phonetic representation within short-term memory. *Clinical Neurophysiology*, *113*, 495–506.

Desjardins, R. N., & Werker, J. F. (2004). Is the integration of heard and seen speech mandatory for infants? Developmental Psychobiology, 45(4), 187–203.

Dopjans, L., Wallraven, C., & Bulthoff, H. H. (2009). Cross-modal transfer in visual and haptic face recognition. *IEEE Transactions on Haptics*, *2*(4), 236–240.

Dupont, S., Aubin, J., & Menard, L. (2005). The study of the McGurk effect in 4 and 5-year-old French Canadian children. *ZAS Papers in Linguistics*, *40*, 1–17.

Fowler, C. A. (1986). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics*, *14*, 3–28.

Fowler, C. A. (2004). Speech as a supramodal or amodal phenomenon. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *The handbook of multisensory processing* (pp. 189–202). Cambridge, MA: MIT Press.

Fowler, C. A., & Dekle, D. J. (1991). Listening with eye and hand: Cross-modal contributions to speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(3), 816–828.

Fuster-Duran, A. (1996). Perception of conflicting audio- visual speech: an examination across Spanish and German. In D. G. Stork & M. E. Hennecke (Eds.), *Speechreading by humans and machines* (pp. 135–143). New York: Springer-Verlag.

Gick, B., & Derrick, D. (2009). Aero-tactile integration in speech perception. *Nature*, *426*, 502–504.

Green, K. P. (1998). The use of auditory and visual information during phonetic processing: Implications for theories of speech perception. In R. Campbell & B. Dodd (Eds.), *Hearing by eye II: Advances in the psychology of speechreading and audiovisual speech* (pp. 3–25). Hove, UK: Psychology Press.

Green, K. P., Kuhl, P. K., Meltzoff, A. N., & Stevens, E. B. (1991). Integrating speech information across talkers, gender, and sensory modality: Female faces and male voices in the McGurk effect. *Perception & Psychophysics*, *50*(6), 524–536.

Green, K. P., & Miller, J. L. (1985).On the role of visual rate information in phonetic perception. *Perception & Psychophysics*, *38*, 269–76.

Harris, C. S. (1965). Perceptual adaptation to inverted, reversed, and displaced vision. *Psychological Review*, *72*(6), 419–444.

Hay, J. C., Pick, H. L., & Ikeda, K. (1965). Visual captur produced by prism spectacles. *Psychonomic Science*, *2*(8), 215–216.

Holmes, N. P., & Spence, C. (2005). Visual bias of unseen hand position with a mirror: Spatial and temporal factors. *Experimental Brain Research*, *166*(3–4), 489–497.

Ito, T., Tiede, M., & Ostry, D. J. (2009). Somatosensory function in speech perception. *Proceedings of the National Academy of Sciences USA*, *106*(4), 1245–1248.

Jerger, S., Damian, M. F., Spence, M. J., Tye-Murray, N., & Abdi, H. (2009). Developmental shifts in children's sensitivity to visual speech: A new multimodal picture-word task. *Journal of Experimental Child Psychology*, *102*, 40–59.

Jiang, J., & Bernstein, L. E. (2011). Psychophysics of the McGurk and other audiovisual speech integration effects. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(4), 1193–1209.

Kayser, C., Petkov, C. I., & Logothetis, N. K. (2008). Visual modulation of neurons in auditory cortex. *Cerebral Cortex*, *18*(7), 1560–1574.

Kinney, J. A., & Luria, S. M. (1970). Conflicting visual and tactile-kinesthetic stimulation. *Attention*, *Perception*, *& Psychophysics*, *8*(3), 189–192.

Konkle, T., Wang, Q., Hayward, V., & Moore, C. I. (2009). Motion aftereffects transfer between touch and vision. *Current Biology*, *19*, 1–6.

Lateiner, J. E., & Sainburg, R. L. (2003). Differential contributions of vision and proprioception to movement accuracy. *Experimental Brain Research*, *151*, 446–454.

Lee, D. N., & Aronson, E. (1974). Visual proprioceptive control of standing in human infants. *Attention*, *Perception*, *& Psychophysics*, *15*(3), 529–532.

Lee, D. N., & Lishman, J. R. (1977). Vision—The most efficient source of proprioceptive information for balance control. *Agressologie*, *18*, 83–94.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*, 1–36.

Macaluso, E. (2006). Multisensory processing in sensory-specific cortical areas. *The Neuroscientist*, *12*, 327–338.

MacDonald, J., & McGurk, H. (1978). Visual influences on speech perception processes. *Perception & Psychophysics*, *24*(3), 253–257.

Massaro, D. W. (1998). Perceiving talking faces: From speech perception to a behavioral principle. Cambridge, MA: MIT Press.

Massaro, D. W., & Cohen, M. M. (1993). Perceiving asynchronous bimodal speech in consonant-vowel and vowel syllables. *Speech Communication*, *13*, 127–134.

Massaro, D. W., Cohen, M. M., Gesi, A., Heredia, R., & Tsuzaki, M. (1993). Bimodal speech perception: An examination across languages. *Journal of Phonetics*, *21*, 445–478.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746–748.

Merabet, L. B., Hamilton, R., Schlaug, G., Swisher, J. D., Kiriakopoulos, E. T., Pitskel, N. B., … Pascual-Leone, A. (2008). Rapid and reversible recruitment of early visual cortex for touch. *PLoS ONE*, *3*(8), e3046. doi:10.1371/jounral.pone.0003046

Mergner, T., & Rosemeier, T. (1998). Interaction of vestibular, somatosensory, and visual signals for postural control and motion perception under terrestrial and microgravity conditions—A conceptual model. *Brain Research Reviews*, *28*, 118–135.

Mottonen, R., Krause, C. M., Tiipana, K., & Sams, M. (2002). Processing of changes in visual speech in the human auditory cortex. *Cognitive Brain Research*, *13*, 417–425.

Munhall, K. G., Gribble, P., Sacco, L., & Ward, M. (1996). Temporal constraints on the McGurk effect. *Perception & Psychophysics*, *58*(3), 351–362.

Musacchia, G., Sams, M., Nicol, T., & Kraus, N. (2005). Seeing speech affects acoustic information processing in the human brainstem. *Experimental Brain Research*, *168*, 1–10.

Norman, J. F., Clayton, A. M., Norman, H. F., & Crabtree, C. E. (2008). Learning to perceive differences in solid shape through vision and touch. *Perception*, *37*, 185–196.

Pekkola, J., Ojanen, V., Autti, T., Jaaskelainen, I. P., Mottonen, R., Tarkiainen, A., & Sams, M. (2005). Primary auditory cortex activation by visual speech: An fMRI study at 3T. *Auditory and Vestibular Systems*, *16*(2), 125–128.

Richler, M. P., Hove, M., & Munhall, K. G. (2003). Gaze behavior in audiovisual speech perception: The influence of ocular fixation on the McGurk effect. *Perception & Psychophysics*, *65*(4), 553–567.

Rosenblum, L. D. (2005). Primacy of multimodal speech perception. In D. Pisoni & R. Remez (Eds.), *Handbook of speech perception* (pp. 51–78). Malden, MA: Blackwell.

Rosenblum, L. D. (2008). Speech perception as a multimodal phenomenon. *Current Directions in Psychological Science*, *17*(6), 405–409.

Rosenblum, L. D., Dias, J. W., & Dorsi, J. (2016). The supramodal brain: implications for auditory perception. *Journal of Cognitive Psychology, 28,* 1–23.

Rosenblum, L. D., & Saldana, H. M. (1992). Discrimination tests of visually influenced syllables. *Perception & Psychophysics*, *52*(4), 461–473.

Rosenblum, L. D., & Saldana, H. M. (1996). An audiovisual test of kinematic primitives for visual speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(2), 318–331.

Rosenblum, L. D., Schmuckler, M. A., & Johnson, J. A. (1997). The McGurk effect in infants. *Perception & Psychophysics*, *59*(3), 347–357.

Rosenblum, L. D., Yakel, D. A., & Green, K. G. (2000).Face and mouth inversion effects on visual and audiovisual speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 806–19.

Saldaña, H. M., & Rosenblum, L. D. (1993). Visual influences on auditory pluck and bow judgments. *Perception & Psychophysics*, *54*, 406–416.

Sams, M., Aulanko, R., Hamalainen, M., Hari, R., Lounasmaa, O. V., Lu, S., & Simola, J. (1991). Seeing speech: Visual information from lip movements modifies activity in the human auditory cortex. *Neuroscience Letters*, *127*, 141–145.

Sams, M., Manninen, P., Surakka, V., Helin, P., & Katto, R. (1998). McGurk effect in Finnish syllables, isolated words, and words in sentences: Effects of word meaning and sentence context. *Speech Communication*, *26*, 75–87.

Sams, M., Mottonen, R., & Sihvonen, T. (2005). Seeing and hearing others and oneself talk. *Cognitive Brain Research*, *23*, 429–435.

Schutz M., & Lipscomb, S. (2007). Hearing gestures, seeing music: Vision influences perceived tone duration. *Perception*, *36*, 888–897.

Sekiyama, K. (1998). Face or voice? Determinant of compellingness to the McGurk effect. In *Proceedings of AVSP'98* (pp. 33–36). Synday: n.p.

Sekiyama, K., & Tohkura, Y. (1991). McGurk effect in non-English listeners: Few visual effects for Japanese subjects hearing Japanese syllables of high auditory intelligibility. *Journal of the Acoustical Society of America*, *90*(4), 1797–1805.

Sekiyama, K., & Tohkura, Y. (1993). Inter-language differences in the influence of visual cues in speech perception. *Journal of Phonetics*, *21*(4), 427–444.

Sekuler, R., Sekuler, A. B., & Lau, R. (1997). Sound alters visual motion perception. *Nature*, *385*, 308.

Shams, L., Iwaki, S., Chawla, A., & Bhattacharya, J. (2005). Early modulation of visual cortex by sound: An MEG study. *Neuroscience Letters*, *378*, 76–81.

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, *14*, 147–152.

Shams, L., Kamitani, Y., Thompson, S., & Shimojo, S. (2001). Sound alters visual evoked potentials in humans. *NeuroReport*, *12*, 3849–3852.

Shams, L., & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in Cognitive Science*, *12*(11), 411–417.

Shimojo, S., & Shams, L. (2001). Sensory modalities are not separate modalities: Plasticity and interactions. *Current Opinion in Neurobiology*, *11*(4), 505–509.

Slater, M., Perez-Marcos, D., Ehrsson, H. H., & Sanchez-Vives, M. V. (2008). Towards a digital body: The virtual arm illusion. *Frontiers in Human Neuroscience*, *2*, 6. doi:10.3389/neuro.09.006.2008

Soto-Faraco, S., & Alsius, A. (2009). Deconstructing the McGurk-MacDonald illusion. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(2), 580–587.

Stekelenburg, J. J., & Vroomen, J. (2007). Neural correlates of multisensory integration of ecologically valid audiovisual events. *Journal of Cognitive Neuroscience*, *19*(12), 1964–1973.

Summerfield, Q. (1987). Some preliminaries to a comprehensive account of audio-visual speech perception. In I. Barbara & C. Ruth (Eds.), *Hearing by eye: The psychology of lip-reading* (pp. 3–51). Hillsdale, NJ: Lawrence Erlbaum.

Summerfield, Q., & McGrath, M. (1984). Detection and resolution of audio-visual incompatibility in the perception of vowels. *The Quarterly Journal of Experimental Psychology*, *36*(A), 51–74.

Thomas, S. M., & Jordan, T. R. (2002). Determining the influence of Gaussian blurring on inversion effects with talking faces. *Attention, Perception, & Psychophysics*, *64*(6), 932–944.

Vatikiotis-Bateson, E., Eigsti, I., Yano, S., & Munhall, K. G. (1998). Eye movement of perceivers during audiovisual speech perception. *Perception & Psychophysics*, *60*(6), 926–940.

Walker-Andrews, A. S. (1997). Infants' perception of expressive behaviors: Differentiation of mutlimodal information. *Psychological Bulletin*, *121*(3), 437–456.

Walker, J. T., & Scott, K. J. (1981). Auditory-visual conflicts in the perceived duration of lights, tones, and gaps. *Journal of Experimental Psychology*, *7*(6), 1327–1339.

Walker, S., Bruce, V., & O'Malley, C. (1995). Facial identity and facial speech processing: Familiar faces and voices in the McGurk effect. *Perception & Psychophysics*, *57*(8), 1124–1133.

Walton, G. E., & Bower, T. G. R. (1993). Amodal representation of speech in infants. *Infant Behavior & Development*, *16*, 233–243.

Watkins, S., Shams, L., Tanaka, S., Haynes, J.-D., & Rees, G. (2006). Sound alters activity in human V1 in association with illusory visual perception. *NeuroImage*, *31*(3), 1247–1256.

How to Use the Website

Purchasers of the print version of *The Oxford Compendium of Visual Illusions* are also entitled to free access to the companion website. It can be accessed at: www.oup.com/visualillusions

Username: Illusoryexperience
Password: Noexperiencenecessary

Please note the capital letters in the log-in credentials. We intend to maintain the website for the lifetime of the edition.

For the most comfortable reading experience, wherever possible, we have provided the illusions as still images. However some effects can only be perceived with motion. We built the companion website with the aim of hosting the accompanying videos and interactive figures. That done, we went further: there is a navigable catalog of every illusion discussed in this compendium.

The website works as both an accompaniment to the book and as a standalone resource with regular updates highlighting some of the most thought-provoking illusions. We hope you enjoy it.

Index

Adaptation
  to apparent motion, 750, 750*f*
  brightness illusion, 379
  contour, 748–49, 748–49*f*
  contrast, pattern specificity of, 746–48, 747*f*
  critique, 154
  dark, 154
  to face aftereffects, 655–57, 656*f*
  flicker, 750–51, 751*f*
  graph-paper effect, 738–42, 741*f*
  jogging aftereffects, 749–50
  luminance profiles, 746, 747*f*
  motion aftereffects, 589–90
  motion sensors, 588–89
  motion standstill illusion, 571
  pattern, aftereffects, 738–42, 741*f*
  ramp aftereffects, 745–46, 745*f*
  sawtooth, 745–46, 745–47*f*
  sensory-motor, 154
  visual, in binocular rivalry, 722–23
Addams, R., 18–19, 586
Adelson, E. H., 369, 533
Afterimages. *see also* motion aftereffects (MAEs)
  boundaries in, 114–15
  breathing light illusion, 376, 376*f*
  facial, 655–56*f*, 655–58
  filling-in illusion, 450–53, 451–52*f*, 453n3
  illusions critique, 150–51
  luminance and color, 323, 323–24*f*
Aglioti, S., 74–77
AGM. *see* attention-generated apparent motion
Agostini, T., 352, 374, 696
Ahluwalia, A., 49–50
Ahumada, A. J. Jr., 467
Akimel O'odham (Pima) tribe, 412*f*, 413
Alais, D., 690
Alakbarov, R., 395, 396*f*
Alam, S., 696
Alberti, L. B., 6
Alberti's window, 6, 7*f*
Albright, T. D., 468, 471, 480
Alcohol, 616
Al-Haytham, I., 155n5
Allegory of Prudence, 610
Alternating brightness star, 407–13, 408–12*f*
Ambiguous apparent motion, 475–77, 475–77*f*
Ambiguous figures in motion
  animal-behavior studies, 663*f*, 664–65
  feature processing, 665–66
  heading effect, 663–65, 663*f*
  motion effects on, 663, 663*f*
  object recognition, 664, 665
  parameters, 664
  shape/motion integration, 665
  2D *vs.* 3D form perception, 663–64
  variations, 664
Ames room
  bistability, 181, 181*f*
  butterfly illusion, 180, 180*f*, 715
  cognition and, 177, 715
  convex/concave parts, 178, 178–79*f*, 181
  critique of, 145–46, 146*f*, 148, 153, 154, 155n2
  curved window, 179*f*, 180–81
  de Heer's sombrero, 180–81, 182–83*f*
  depth inversion, 174, 177, 177*f*
  eye position movement, 177, 177*f*
  fence, 180, 181*f*
  fundamental element in, 174, 174*f*
  illusory-viewed orientation, 174, 174*f*
  motion parallax, stereoscopic disparity, 281
  rotating window, stationary observer, 174–75, 175*f*, 177
  shape distortion, perspective-based, 175–77, 176*f*
  stationary window, mobile observer, 175, 175*f*
  stingray illusion, 178–80, 179–80*f*
  3D objects, 180, 181*f*
  tilted window, 181–83, 182*f*
  variants of, 177–78
  windsurfer illusion, 176, 176*f*
Amodal completion
  angle interpolation, 675–76, 675*f*
  approximation *vs.,* 676
  attentional focus, 673
  of boundaries, 90–91, 91*f*, 93, 99
  breathing light illusion, 374, 375*f*
  contour curvature polarity, 676
  described, 674–75
  distorted hexagon, 673, 674*f*
  distortion, occlusion, 673–74
  eccentric square, 674*f*, 675, 675*f*
  field model, 676
  Gerbino illusion, 673, 674*f*, 680n1
  global shape, 676
  good continuation, 674–75
  human face, body illusions, 255–56, 256*f*
  identity hypothesis, 674
  as mental representations, 127–30, 132*f*
  minimal path, 674–76
  modal *vs.,* 346, 346*f*, 347
  orientation, 676
  phenomenal rounding/trajectory, 675–76
  Poggendorff illusion, 673, 674*f*
  scale dependence, 676
  support ratio, 676
  T-junctions, 673, 674*f*, 679–80, 680*f*
  weak identity hypothesis, 674–75
Anderson, B. L., 328, 443, 682, 683*f*
Ando, H., 784
Ando, S., 626
Annan, V., 354*f*
Anstis, S. M., 149, 323, 325, 376, 438, 467, 468, 471, 531–33, 588, 594, 750, 752
Antigravity slope illusion
  explanation of, 295–97, 296–97*f*
  magnet-like slopes, 147–48, 147*f*, 295–97, 296*f*
  principles, construction of, 295, 296*f*
  road, 296–97, 296–97*f*
Aperture capture illusion
  described, 700, 701*f*, 703–4
  object tracking, 700, 701*f*, 702, 704
  parameters, 703, 703*f*
  perceived *vs.* true velocity, 700, 701*f*
  persistence, 702, 703
  position updating, 702, 703
  relatability principle, 702, 702*f*
  research applications, 700–2
  retinal painting, 701*f*, 704
  spatiotemporal relatability, 701*f*, 702–4
  variants, 703, 703*f*
Aperture problem
  across multiple apertures, 559–60, 559*f*
  barber pole illusion, 557–59, 558*f*
  boundaries in, 106–7, 559–60, 559*f*
  conjunctive sensitivity, 556–57
  described, 556–57, 557*f*

Aperture problem (*cont.*)
  discontinuity hypothesis, 559–60, 559*f*
  in motion illusions, 81, 81*f*
  object recognition, 559*f*, 561
  receptive field size, 556, 557, 561–62
  rotating tilted lines illusion, 491
  spinning ellipse speed illusion, 170, 170*f*
  visual system segmentation, integration, 560–61
Apparent motion
  adaptation to, 750, 750*f*
  attention-generated (*see* attention-generated apparent motion)
  compensatory prediction, 565–66
  described, 563, 564*f*
  filling-in, 566
  Fitts's law, 567
  of human body, 564–66, 565*f*, 567, 567*f*
  parameters, 566, 567*f*
  perception-action coupling (embodied perception), 566
  shortest path constraint, 563–65
  stimulus onset asynchrony (SOA), 564–65, 565*f*, 567, 567*f*
Arabzadeh, E., 328
Arcimboldo, G., 633, 635*f*
Arend, L. E., 444
Aristotle, 5, 6, 8, 18
Artal, P., 759
ART models, 112–14, 113–14*f*
ARTSCAN model, 114
Ashida, H., 86, 87, 502, 507*f*, 510, 586, 589, 784
Astigmatism, 758–59, 759*f*
Attention-generated apparent motion
  depth/texture gratings, 527, 528*f*
  described, 527, 528*f*
  feature-salience motion perception, 527
  full-wave/color gratings, 527, 528*f*
  full-wave/half-wave gratings, 527, 528*f*
  process models, 529, 529*f*
Attneave, Fred, 412*f*, 413
Aubert, H., 22, 305
Autism, 636, 657
Autokinetic effect, 21–22
Automatic centroid extraction model, 161–63, 162*f*
Autostereograms, 779
Axelrod, V., 630
Aydin, M., 688

Baar, J., 55
Babadi, B., 328
Baboons, 636
Backscroll illusion
  brain areas activated by, 487–88
  center-surround antagonism, 487
  described, 486, 487*f*
  motion-nulling, 486
  object motion recognition, 486–87
  retinal motion signals, 488
  signal dots stimulus, 486–87
  strength quantification, 486
Backus, B. T., 509
Bacon, F., 7
Banks, M. S., 239
Barber pole illusion, 106, 489, 557–59, 558*f*
Bar-cross-ellipse illusion
  bars oscillating in depth, 165, 167*f*
  breathing square, 167–68, 168*f*
  fat apertures, 167, 168*f*
  morphing cross, 165, 166*f*
  narrow apertures, 167, 167–68*f*
  neural mechanisms underlying, 168–69
  occluded ellipse, 165–66, 166*f*
  overview, 165, 166*f*
  rectangle, 168, 168*f*
  rotating elliptical aperture, 166, 167*f*
  skinny ellipse, 167, 167*f*
  variants of, 166–68, 167*f*
Barlow, H. B., 413, 586
Barton, J. J., 657
Bathtub illusion
  binocular viewing, 239
  edges, 238–39, 239*f*
  flatness cues, 239–40
  mirror symmetry, 238–39, 239*f*
  objects *vs.* pictures projections, 238–40, 239*f*
  retinal projection as guess, 238
  vantage point changes, 238–40, 239*f*
Baylis, G., 625
Bayne, K. A. L., 56
Beaunis cubes, 11–12, 11*f*
Beck, J., 101, 384
Beer, A. L., 510
Beghi, L., 597
Benary, W., 353
Benhar, E., 55–56
Bentley, A. M., 48–49
Bergen, J. R., 362, 401, 402*f*, 533
Bergström, S. S., 369, 394
Berkeley, G., 6
Bernstein, L. J., 663, 664
Berry, J. W., 43
Berry, M., 288n10
Bertamini, M., 609, 610, 611
Bertin, E., 636
"Between Earth and Sky," 397*f*
Beyil, J., 413
Bhatt, R. S., 636
Bicycle spokes illusion, 469, 469*f*
Biederlack, J., 353
Biederman, I., 412*f*, 413
Billino, J., 490
Binding by synchrony hypothesis, 352–53
Bindman, D. R., 671
Binocular rivalry
  continuous flash suppression, 721, 723*f*
  described, 721, 722*f*, 728n1
  dichoptic displays, 721, 722*f*
  dominance, 723
  Mondrian patterns, 721, 723*f*
  motion aftereffects, 722
  neural responses, 723
  overriding, 721
  perception bias, 721–22
  research applications, 723
  stereopsis and, 723
  subliminal perception, 722
  visual adaptation, 722–23
Binocular vision. *see* stereopsis
Blackshot mechanism, 670–71, 671*f*
Blake, R., 488
Blakemore, M. R., 470
Blakeslee, B., 328, 329, 391, 415–20, 416*f*, 418*f*, 444
Blanz, V., 655
Blaser, E., 529
Blind spot, 94–95, 95*f*
Bloj, M. G., 383, 394
The Blue Boy, 324–25, 324–25*f*
Blur adaptation, induction
  adaptive optics in, 759–60
  aftereffects, 756–57, 757*f*
  in complex images, 758, 758*f*
  described, 756, 757*f*
  orientation-selective, 758–59, 759*f*
  perceived focus, 756, 757*f*, 758
  perceptual *vs.* neural norm, 759–60
  renormalization aftereffect, 758
  spatial frequency, 756–58, 757*f*
Body illusions. *see* human face, body illusions
Bogart effect
  described, 624, 625*f*
  gaze direction detection, 624–25
  gaze-estimation process, 624–26, 626*f*
  geometry-based cues, 626
  luminance-based heuristic, 625–27, 626*f*
  robustness of, 624, 626*f*
Boi, M., 685, 686, 686*f*
Bolton, C., 46
Bolton, R., 46
Bonfiglioli, C., 219
Bonté, M. C., 43–44
Borges, J. L., xxi
Boring, E. G., 12
Born, R. T., 107, 493
Bornstein, M. H., 46
Borschke, A., 19–20
Bosten, J. M., 340, 746
Botti's illusion, 696, 697*f*
Boundaries in visual world. *see also* subjective contours
  afterimages, 114–15
  amodal recognition, 90–91, 91*f*, 93, 99
  aperture problem, 106–7, 559–60, 559*f*
  ART models, 112–14, 113–14*f*
  ARTSCAN model, 114
  attentive matching, 113–14, 114*f*
  barber pole illusion, 106, 489
  Bayesian inference in brain design, 111–12
  bipole grouping, 97–98
  blind spot, 94–95, 95*f*
  collinearity, 101
  complementary processing streams, 98–99, 98*f*, 104–6, 105*f*
  complementary properties, 96*f*, 98
  completion/surface filling-in, 91*f*, 96–98, 96*f*, 150
  context in, 108–9, 109*f*
  contrast polarity, 93–94, 93–94*f*, 96, 98
  Dalmatian in snow illusion, 100–101, 100*f*
  dimensionality, 91–92, 91*f*, 102–4, 104*f*

directional filtering, 108–9, 109*f*
distance, relative orientation, 101–2, 101*f*, 103*f*
edges as illusion, 96–97
Ehrenstein illusion, 91*f*, 93
end gaps, 97
expectations, imagination, 90
FACADE theory, 92, 99, 102, 112, 114
feature tracking, 106–7
FORMOTION model, 104, 107
illusory contours, apparent motion of, 104–5, 105*f*
invisibility, 93–94, 93–94*f*, 101–2
Kanizsa square, 91–92, 91*f*, 94, 94*f*, 96, 99
LAMINART theory, 92, 102, 112–13, 113*f*
line ends as illusion, 97
MODE model, 107–8, 108*f*, 110–11, 111*f*
motion capture, 107–8, 108–9*f*
neon color spreading, 97–100, 97*f*, 99*f*
noise-saturation dilemma, 110
object occlusion, 95–96
offset grating, 90, 91*f*, 97
parallelopipeds, 102–4, 104*f*
perceptual decision-making, 107–8, 108*f*
prior knowledge/perception relationships, 91*f*, 92
probability distribution, 110–11, 110–11*f*
smooth pursuit eye movements, 105–6
stability-plasticity dilemma, 112–13, 113*f*
stratification images, 91–92, 91*f*
surface-shroud resonance, 114
surfaces in art, movies, 91–93, 91*f*, 93*f*
symmetry-breaking, 99
texture grouping, 100–102, 100–103*f*
uncertainty, hierarchical resolution of, 95, 97
unconscious inferences, 92
visual processes, 92–93
watercolor illusion, 99–100, 99*f*, 434, 434*f*
Bours, R. J. E., 467
Bowditch, H. P., 19, 586
Brady, T. F., 766
Breathing light illusion
afterimages, 376, 376*f*
albedo hypothesis, 374, 375*f*
amodal completion, 374, 375*f*
boundaries, 373–76
brightness variation, 374
color perception, 373
described, 373–74, 374*f*
luminance gradient, 374–76, 375*f*
size variation, 374
Breitmeyer, B., 689
Brentano, F., 159, 300
Bressan, P., 334, 354*f*, 369
Brewster, D., 8, 28, 614
Brief window hypothesis, 578–79, 579–80*f*
Brightness illusion
adaptation, 379
contrast gain, 380
described, 379, 380*f*
theories of, 379–80
visual competition, 380–81
Brislin, R. W., 46–47, 49
Brooks, R., 124
Bross, M., 305
Brown, R., 709
Brown, T., 636
Bruce, V., 616, 636
Brunelleschi, F., 6
Bulatov, A., 306
Bullseye illusion, 446
Bulthoff, H. H., 75
Burgh, P., 374
Burke, L., 674
Burton, G., 263
Butterfly illusion, 180, 180*f*, 715

Café wall illusion, 119, 120*f*
Caldara, R., 764
Calder, A. J., 636, 643
Cambridge Expedition. *see* cross cultural studies
Camera obscura, 6
Cannabis, 616
Cannon, M. W., 340
Cantor, C. R. L., 510
Cap or cup illusion
augmented shading cue, 227*f*, 229–31, 230–31*f*
conical curves, 231, 232*f*
dynamic full-wave stimuli, 405–6
ecological considerations, 227–28, 228*f*
edges, 231, 232*f*
ellipsoids, 231, 232*f*
illumination, 229–30, 230*f*
inverse optics, 227*f*, 228
invisibility, 230–32, 231*f*
invisible saddle, 227, 227*f*
luminance modulation function, 404, 404–5*f*
outline change, 230–31, 230–31*f*
probability density function, 228, 228–29*f*
quadric shapes, 227, 228*f*
shading cue, 227*f*, 228–29, 229–30*f*
texture, contrast in, 404–5, 404–5*f*
Caputo, G., 458n3
Carey, S., 633
Carlin, J. D., 643
Carlin, P., 341
Carlson, J. A., 46
Carr, H. A., 22
Carrasco, M., 150, 381, 751
Casco, C., 290
Castiello, U., 219
Cast shadows. *see* shadows
Catching patches illusion, 449–50, 450–51*f*, 453nn1–2
Catharine of Sienna (Saint), 136n2
Cavanagh, P., 467, 468*f*, 589, 690
Celestial illusion. *see* moon (celestial) illusion
CGT model, 360
Chabris, C. F., 716*f*
Challinor, K., 532, 533
Chan, W. M., 681*f*
Change blindness. *see* silencing
Charles, J. P., 438
Checkerboards, 292, 292*f*
Chen, L., 688, 688*f*
Chen, V. J., 346, 347
Chevreul, M. E., 29, 30*f*, 356
Chevreul illusion (luminance ramp)
CGT model, 360
described, 356, 356–57*f*, 360n1, 404, 404*f*
difference of Gaussians (DoG) model, 357–60, 359*f*
edges, 359–60
lateral inhibition, 356–60, 356*f*
luminance modulation function, 404, 404–5*f*
variants, effects of, 357–58, 358*f*
Chey, J., 107, 493
Chilcoat, M., 595n1
Chinese lantern illusion
banding, 438–40, 438–40*f*
border distortion, 437–39, 438*f*
contour deviation, 438–40, 438–40*f*
CP rule, 439–40, 439*f*
described, 437, 437*f*
luminance gradient, 438, 438*f*, 439
path of conjunction, 438*f*, 439–40
region of equiluminance, 438, 438*f*
Choi, P. P. C., 50
Chopstick illusion, 479–82, 481–82*f*
Chromatic Mach card
Bayesian model, 384, 385*f*
brightness gradients, 384
brightness perception, 382, 382*f*
color constancy, 382–83, 386
described, 383, 383*f*, 394
mutual illumination, 384–85
object shape, color, illumination, 383–85, 383*f*, 385*f*
parameters, 385–86, 386*f*
Chromatic shadow illusions, 396–98, 397–98*f*
Chubb, C., 341–42, 446, 670, 671
Cicerone, C. M., 344, 345, 346, 347
Cline, M. G., 624
Cloete, S., 501
Cohen, M. A., 360
Cohen, R., 571
Color assimilation
causes of, 388–90
chromatic White's effect, 389*f*, 391
contrast normalization, 389*f*, 390, 391
contrast *vs.,* 388
layer decomposition, 390, 391
neon color spreading, 390*f*, 392
neural blurring, 388–90, 389–90*f*
overview, 388, 389*f*
ring patterns, 390*f*, 391
S-cone responses, 390*f*, 391
terminology, 388
T-junctions, 328, 391
Color dove illusion
afterimage effect, 753*f*, 754
described, 752, 753*f*
filling-in, 752, 753*f*, 754
flickering, 752, 753*f*
parameters, 752, 753*f*
as positive aftereffect, 752, 753*f*
simultaneous contrast, 752–54, 754*f*
variants, 752, 753*f*

Color from motion
binocular combination in, 346–47
completion, modal *vs.* amodal, 346, 346*f*, 347
contour formation, 345–46
depth, 347
described, 344, 345*f*
deuteranope, 345–46
distinctive features of, 347
functional significance of, 347–48
luminance contrast, 345–46
luminance/visual scene organization, 346
perception of motion in, 344–45, 345*f*
regulation of, 346–47
transparency, perception of, 347
Color sphere, 121, 122*f*
Complementary processing streams, 98–99, 98*f*, 104–6, 105*f*
Compression gratings, 422, 423*f*
Configural shape illusion, 269, 271–72, 271*f*
Contrast, assimilation. *see also* size contrast, assimilation
contrast polarity, 93–94, 93–94*f*, 96, 98
critique, 149–51
described, 65, 65*f*
Ebbinghaus-Titchener illusion, 65, 65*f*
geometrical optical illusions, 65, 65*f*
Hering illusion, 150
luminance (*see* luminance contrast polarity)
Müller-Lyer illusion, 65, 65*f*, 150
nonhuman species studies, 58, 58*f*, 60, 60*f*, 61
Orbison illusion, 150
Poggendorff illusion, 65, 65*f*
Ponzo illusion, 65, 65*f*
simultaneous effects, 150
size (*see* size contrast, assimilation)
wedding cake illusion, 442–46, 443*f*
Zöllner illusion, 65, 65*f*, 150
Contrast asynchrony
coarse *vs.* fine-scale information, 772
color models, 770–72, 771*f*
described, 767–68, 767–68*f*
perceptual organization, 769–70, 770*f*
single *vs.* two fields, 770–72, 771*f*
thin edges, 768–69, 768–69*f*
Contrast contrast illusion
color and, 338*f*, 342
contrast polarity selection, 338*f*, 341–42, 341*f*
described, 337, 338*f*
ecological accounts of, 338*f*, 342
individual differences, 340–41
schizophrenia, schizoaffective disorder, 341
surround-induced enhancement, 340–41
textural granularity, 338–40*f*, 340
texture contrast sensing, 337–42, 338–39*f*, 341*f*
Conway, N., 468
Cook, M., 490
Cook, N. D., 281
Cooper, L. A., 663, 664
Coordinate frames, 513–14, 514*f*
Coren, S., 46, 51, 264, 265
Cornell, A., 716
Cornsweet, T. N., 351
Cornsweet illusion, 373, 404, 405*f*
Correani, A., 456
Corrow, S., 616
Cosmetics, 622
Couple with Mirror, 610
Craik-O'Brian-Cornsweet illusion, 130
Creem, S. H., 187, 192
Crick, F., 383
Cross cultural studies
Africans, 43–51
age, 42, 44, 46, 49–50
Americas, 46–49
box bridge illusion, 38–40, 39*f*
calipers, 48–49
carpenteredness, 40–46, 48–49, 51
children, 42, 44, 46, 49–50
color alternation, 41–42, 41*f*, 45–46
consonance, 47
depth cues, 46–48
Ebbinghaus (Titchener) circles, 47, 51
ecological hypothesis, 41, 45
eidolic effect, 48
epitomic effect, 48, 48*f*
Eskimos, 43
Fang, 43
foreshortening, 41, 42, 44, 47, 51
gender differences, 43, 49–50
Hermannsburg, 42, 43
holistic stimuli, 51
horizontal–vertical illusion, 41–47, 41*f*, 50, 51
hypotheses, 40–43
Ijaw, 43
Kalahari Bushmen, 43
methodology, 40–42, 41*f*
ML/dichosection scores, 45
Müller-Lyer illusion, 38–43, 39*f*, 41*f*, 44–46, 49–51
Navajos, 44
Nuristanis, 44
Oppel-Kundt figure, 47–48, 50
perspective drawing, 41–42, 41*f*, 44, 46–48
pictorial perception, 46–49
pigmentation hypothesis, 45–46
Poggendorff illusion, 40–42, 41*f*, 44–45
Ponzo illusion, 46–47, 49–50
range of illusions, 44–45
reversals, 48
Sander parallelogram, 41–44, 42*f*, 46, 49–51
Sholagas, 40
surface-grain effect, 44
susceptibility, developmental changes in, 50
susceptibility to illusion, 43–51
Toda, 40, 44, 48
Torres Straits study, 38–40, 39*f*
Uralis, 40
Western Europeans, 40, 43–51
Yuendumu, 43
Zöllner illusion, 44–45
Crossover motion, 467, 468*f*
Crosstalk theory, 69–70, 70*f*
Crussaire, Pierre, 8, 9, 9*f*, 11
Csathó, Á., 293, 438
Culham, J., 589
Cumming, B. G., 104
Curved grid non-illusions
Baumgartner model, 362–64, 363*f*
collinearity, 364–65
curvature, 363, 363–64*f*
distortion tolerance, 364–65, 364–65*f*
edge segments, 364–66, 364*f*
Hermann grid, 362–66, 363–64*f*
lateral inhibition, 362–64, 363*f*
Lingelbach's scintillating grid, 362, 363*f*, 364–66
sinusoid grids, 363, 363–64*f*
spots, 363–65, 363*f*, 366
Cusanus, N., 639, 649n2
Cut-The-Knot, 673

Dakin, S., 341
Dallenbach, K. M., 22
Dalmatian in snow illusion, 100–101, 100*f*
D'Arcy, Chevalier Patrice, 14
Dark adaptation, 154
Darwin, E., 20–21
Darwin, R., 15, 20
Davidoff, J., 51
Davis, C. M., 46
Davis, R. T., 56
Daw, N., 323
Dawson, J. L. M., 50
Day, R., 517
"Daydreaming," 277, 280*f*
Dean, M., 636
DeAngelis, G. C., 104
DeCarlo, D., 617
de Heer circle, 176*f*, 180–81, 182–83*f*
Dekle, D. J., 788
De Lafuente, V., 401
Delboeuf illusion
dynamic illusory size contrast, 260–61, 260–61*f*
eye makeup, 252, 253*f*
inducer conditions, 263
information uncertainty hypothesis, 259–60, 259*f*
length, 206, 206*f*
numerosity, 696, 697*f*
pool and store model, 264–65
size contrast, assimilation, 262–63, 262*f*
DeLucia, P. R., 71, 185, 187–88
De Marchi, S., 696
Depth inversion illusions, 275, 614. *see also* hollow mask illusion; reverspectives
Deręgowski, J. B., 48–49, 50, 234
Derrick, D., 788
Desaguliers, J. T., 6
Descartes, R., 6, 775, 776*f*
DeSouza, J. F., 74–77
Diamond, R., 633
Difference of Gaussians (DOG) model
grating induction, 416, 416*f*
Vasarely nested squares, 407, 409*f*, 411*f*
Differential activation theory, 688
Dima, D., 616

Disappearing hand trick, 154
Dissemination of illusions, xxi
Distribution shift model, 586
Dobkins, K. R., 471
Dominguez, K. E., 54–55
Donguri pattern, 84–87, 85–87*f*
Donnelly, N., 636
Dot brightness density modulation, 422–25, 423–24*f*, 425n1
Double-anchoring model of lightness, 334–36
Double neon color illusion, 448–49, 449*f*
Drifting spines illusion, 501
Driver, J., 625
Dücker, G., 55
Duncker, K., 18, 512–14, 516, 520
Dungeon illusion
  chromatic, 334, 334*f*
  countershading, 332, 336n1
  double-anchoring model of lightness, 334–36
  double-decrement, 332–33, 333*f*, 335
  double-increment, 333, 333*f*, 335
  double-increment contrast prediction, 335, 335*f*
  elements of, 332, 333*f*
  environmental effects prediction, 333*f*, 336
  luminance ratios, 332, 333*f*
  luminance similarity, 334
  paradoxical, 333–36, 333*f*
  variants, prediction of, 334*f*, 335
Durgin, F. H., 192, 193, 194, 749
Dürsteler, M. R., 517
Dworkin, L., 305
Dynamic illusory size contrast
  Delbeof, Ebbinghouse, 260–61, 260–61*f*
  effect, principles of, 258–59, 258–59*f*
  information uncertainty hypothesis, 259–60, 259*f*
Dynamic observer hypothesis, 639–40, 640*f*
Dziurawiec, S., 48
D'Zmura, M., 342

Ebbinghaus, H., xx, xxii
Ebbinghaus-Titchener illusion
  contrast, assimilation, 65, 65*f*
  cross-modal matching, 75, 76
  dynamic illusory size contrast, 260–61, 260–61*f*
  inducer conditions, 263
  information uncertainty hypothesis, 259–60, 259*f*
  length, 206, 206*f*
  magnitude of, 263–64
  maximum-aperture finger configuration, 74–77
  nonhuman species studies, 55, 58–59, 59–60*f*, 61
  pool and store model, 264–65
  relative size comparison, 264
  size contrast, assimilation, 263–64, 263*f*, 290
  size scaling, 65*f*, 67–69
  target/inducer separation, 263
  two visual systems hypothesis, 74–77
Econopouly, J., 670
Ehrenstein, W., 18, 93
Ehrenstein illusion, 91*f*, 93, 681, 681*f*
Einav, S., 413
Elevation. *see* pitchroom illusion
El Greco fallacy, 759
Enigma illusion. *see also* rotating tilted lines illusion
  described, 490, 496, 497–98*f*
  line tilt, 496–99, 498*f*
  luminance effects, 499
  T-junctions, 498*f*, 499
  variations of, 496–99, 497–98*f*
Erdfelder, E., 306
Erkelens, C. J., 239
Escalator illusion. *see* Fraser-Wilcox illusion
Escher, M. C., 92, 584
Euclid, xx, 15–16, 512
Euclid's cone, 6, 7*f*
Evans, R. M., 455, 458n3

FACADE theory, 92, 99, 102, 112, 114
Face aftereffects
  adaptation to, 655–57, 656*f*
  age effects, 657
  autistic children, 657
  bias in, 657–58
  described, 655–56, 655–56*f*
  distortion, 655, 656*f*, 657
  exemplar-based models, 656, 657
  expression, 657
  face coding, 656–57, 656*f*
  gender, 657
  identity, 657
  inversion effect, 657
  norm-based models, 656–57, 656*f*
  plasticity in, 656
Face illusions. *see* human face, body illusions
Face processing. *see* Margaret Thatcher illusion; wobbling face illusion
Face recognition. *see* Presidential illusion
Face-vase illusion, 694
Facsimile, 148, 154
Fagot, J., 59
Fahle, M., 75
Falls of Foyers, 18, 19*f*, 20, 31*f*
Fantoni, C., 674–76
Faraday, M., 15, 548
Farah, M. J., 634
Faubert, J., 87, 501
Faul, F., 306
Fazl, A., 328
Fedorovskaya,E., 394
Fick, A., 27
Figueroa, J. G., 150
Filled space illusion. *see* Oppel-Kundt illusion
Filling-in. *see also* subjective contours
  afterimage filling-in illusion, 450–53, 451–52*f*, 453n3
  apparent motion, 566
  boundaries, 91*f*, 96–98, 96*f*, 150
  catching patches illusion, 449–50, 450–51*f*, 453nn1–2
  chromatic (*see* color dove illusion)
  color dove illusion, 752, 753*f*, 754
  double neon color illusion, 448–49, 449*f*
  grating induction, 416–19, 419*f*
  illusions critique, 151
  motion-induced blindness, 710*f*, 712
Fisher, G. H., 188
Fitzgibbon, A. W., 145
Flash grab illusion, 483–84, 483–84*f*
Flicker augmented contrast, 325–26, 325*f*
Floating-motion illusion, 501
Fluttering-heart illusion, 501
Flying bugs illusion, 478–79, 478–79*f*
Foley, J. M., 327, 415
Footsteps illusion, 469–71, 469–71*f*, 501
"Forced into Reverse Perspective," 277, 278, 280*f*
FORMOTION model, 104, 107
Foster, C., 648
Four-stroke motion
  explicit directions role, 544, 544–45*f*
  temporal impulse response (TIR), 533, 533*f*
Fowler, C. A., 788
Fox, C. J., 657
Fraisse, P., 43
Francis, G., 452
Franz, V. H., 75, 77
Fraser, A., 501
Fraser, J., 293
Fraser-Wilcox illusion
  age effects, 505–9, 510*t*
  array interaction studies, 510
  color enhancement, 504–5, 509*f*
  darkening, contrast effects, 501–2, 503*f*
  described, 501, 502*f*
  disambiguation, 501–2, 503*f*
  eye movements model, 509–10
  fMRI studies, 510
  history of, 500–501, 500*f*
  luminance dependency, 502, 505*f*, 507–8*f*, 604
  optimized, 502–4, 507–9*f*, 510
  perceptual dimorphism, 501–2, 503–6*f*
  peripheral *vs.* central vision, 501, 510
  timing differences model, 509, 510
  T-junctions, 501
Freaking News, 648
Freezing rotation illusion
  center-surround relative motion, 517
  critical parameters, 519–20, 519*t*
  described, 516, 516*f*
  fixed constant velocity, 517
  mechanisms, 520–21
  motion capture, 517
  reference frames, 517
  stimulus composition, 516–17, 516*f*, 518*t*
  variants, parameters, 517–18
Freire, A., 254
Frisby, J., 586, 745
Fujii, Y., 252
Fujimoto, K., 486, 487
Fujita, K., 56
Fullenkamp, S. C., 340
Furrow illusion, 472–73, 472–73*f*

Galileo, G., 5*f*, 6
Galmonte, A., 374
Ganel, T., 75, 77
Gaze direction. *see* Bogart effect; Mona Lisa gaze effect

Gaze-estimation process, 624–26, 626*f*
Gegenfurtner, K. R., 75, 490
Geier, J., 362, 365, 403
Geiger, D., 682*f*
Gelade, G., 352
Gelb effect, 395, 396*f*
Geometrical mosaics, 8, 9*f*, 778
Geometrical optical illusions. *see also specific illusions*
  computational theories, 80
  concepts, definitions, 64
  constancy scaling theory, 65*f*, 67–68
  context, 207
  contrast, assimilation, 65, 65*f*
  cross-orientation inhibition, 65–66, 66*f*
  crosstalk theory, 69–70, 70*f*
  empirical theory of vision, 71–72
  functional theories, 67–70, 67–70*f*
  history, 26–27*f*, 26–29, 29*f*
  isoluminance effects on, 206–7*f*, 206–8
  length, 206, 206*f*
  modal completion, 66, 66*f*
  Oppel, 298–302, 299*f*, 301*f*
  orientation, 207
  perceived size, distance, 65*f*, 67–69, 69–70*f*
  perceptual theory role, 64
  physiological theories, 65–67, 66*f*
  scene statistics, 71–72, 71*f*
  transitivity, 64–65, 65*f*
  transversal collinearity, 70–71, 70*f*
  weighted positional averaging, 159–63, 160*f*, 162*f*
Gerbino, W., 674–76
Gerbino illusion, 673, 674*f*, 680n1
Ghost phi, 593–95, 593*f*, 595*f*
Giaschi, D., 750
Gibson, J. J., 68, 148
Gick, B., 788
Gilchrist, A. L., 328, 334, 354*f*, 360, 369, 384, 394, 395, 444
Gillam, B., 69–70, 144, 185, 681*f*, 682*f*
Gilpin, H., 154
Gindy, S. S., 391
Giora, E., 374
Girgus, J. S., 264, 265
Girshick, A. R., 239
Glare effect
  described, 454–55, 454*f*
  dynamic, 454–55, 455*f*
  glare parade, 455–56, 455*f*, 457nn1–2
  luminance gradients, 456
  luminance pseudo-ramps, 457, 458*f*
  luminosity perception, 455–56
  phantom illumination, 456, 456*f*
  ratio-to-white hypothesis, 455
  simultaneous lightness contrast displays, 455–56*f*, 456, 458*f*, 458n3
  sun figure, 456, 457*f*
  superwhite, 455, 457n2
Glaser, D. A., 499
Glennerster, A., 67, 145
Gobell, J., 275
Goethe, J. W., 24, 29
Gombrich, E. H., 612
Goodale, M. A., 74–77, 187, 265
Gori, S., 374, 376, 489, 490, 491–93, 496, 498*f*
Graham, M. E., 147
Granrud, C. E., 616
Graph-paper effect
  characteristics of, 734–37
  described, 733–34, 733–35*f*
  edge polarity, 735–37, 735*f*, 738–39*f*
  flickering lights, 738
  inhibition, disinhibition, 740–42, 741*f*
  luminance profiles, 735, 738*f*
  motion, 734
  orientation, 734–35, 736–37*f*
  pattern adaptation aftereffects, 738–42, 741*f*
  research applications, 737–38
  stereo depth, 737
Grating induction
  brightness/lightness profiles, 415–16, 416*f*
  as distance function, 419–20, 419–20*f*
  DOG filtering model, 416, 416*f*
  edge detectors, 416
  filling-in, 416–19, 419*f*
  ODOG model, 416, 417*f*, 419–20
  spatial filtering, 415
  spatiotemporal surface characterization, 417–19, 418*f*
  strength of, factors affecting, 415
Great circle model, 246–48, 247*f*
Green, M., 595n1
Green, R. T., 69
Greenlee, M. W., 510
Gregor, A. J., 42, 43, 50
Gregory, R. L., 67–68, 136n11, 145, 146, 155n1, 185, 186, 438, 467, 544, 614, 618, 770
Grossberg, S., 92, 100, 103, 104, 106, 107, 110, 360
Grossman, E. D., 488
Guilford, J. P., 22
Guo, K., 208
Gustaffson, K.-A., 394
Gyani, A., 153, 281

Haddon, A. C., 38
Hadwin, J. A., 636
Hagen, D. J., 44
Hall, G. S., 19, 586
Haller, A. von, 8
Hamburger, K., 207, 208, 489, 490, 496, 498*f*, 499, 770
Hancock, P. J. B., 648
Harris, A. V., 55
Harris, J., 17–18
Hartline, H. K., 741, 742
Hautaluoma, J. E., 44
Hay, D. C., 635
He, Z., 584, 687, 688*f*
Heading effect, 663–65, 663*f*
Head size illusion, 254, 255*f*
Heckel, A. H., 510
Hein, E., 687, 688*f*, 690
Heinemann, E. G., 420
Heitger, F., 83, 683
Hellawell, D., 635
Helmholtz, H. von, 22, 24, 25*f*, 31*f*, 92, 113, 145, 221, 286, 288n11, 305, 308, 393, 704
Helmholtz squares illusion, 26*f*, 27
Helmsley, D., 341
Helson, H., 388, 442
Hempstead, L., 675
Henry, J., 752
Herbert, A. M., 87, 501
Hering, K. E. K., 31*f*, 303, 305
Hering's shadow illusion
  contrast, assimilation, 150
  described, 394–95, 395*f*
  length, 206, 206*f*
  in real-world settings, 188
Hermann, L., 28, 31*f*, 362
Hermann-Hering grids
  dark diagonals, 400–401, 401*f*
  described, 362–66, 363–64*f*
  Fourier model, 401
  history, 25, 28, 29*f*, 38–40, 39*f*, 400
  low-pass filtered, 401, 402*f*
  retinal receptive fields, 400–401, 401*f*
Herskovits, M. J., 48
Herzog, M., 685, 686, 686*f*
Hescheles, L., 19–20
Hidden images
  Bayesian inference, 777
  carrier patterns, 775, 775*f*
  depth, 779, 779*f*
  hybrid faces, 775–76*f*, 777
  orientation, symmetry, 778–79, 778–79*f*
  perceptual organization, 777–78, 777*f*
  perceptual portraits, 774–78, 775–78*f*
  spatialized, 777
Higashiyama, A., 195
High phi, 592–93*f*, 592–95, 595*f*, 595n1
Hill, H., 152, 616
Hill, R. M., 586
Hill, W. E., 12
Hine, T. J., 490
Hine illusion, 501
Hisakata, R., 510
History
  ambiguous figures, 8–13, 9–11*f*, 13*f*
  color juxtapositions, 28–30, 30*f*
  contrast, border enhancement, 29–30, 30*f*
  depth fluctuations, 8, 9*f*, 12
  figure–ground alternation, 8–11, 9–10*f*
  flags, heraldic figures, 10*f*, 11
  Hermann-Hering grids, 25, 28, 29*f*, 38–40, 39*f*
  Mach bands, 30–31, 30*f*
  modern period, xx
  moon (celestial) illusion, xix–xx, 5–6, 5*f*, 308
  motion illusions (*see* motion illusions)
  orientation illusions, 26–27, 27*f*
  overview, 3–4, 4*f*
  pictorial referents, 12–13
  pictorial representations, 6–8, 7*f*
  Purkinje shift, 23*f*, 24
  reversals, 8, 9, 9*f*, 12
  Roman mosaics, 8–9, 9–10*f*, 12
  subjective visual phenomena, 22–25, 23*f*, 25*f*
  theoretical interpretations, 31–32
  tiled floors, 11–12, 11*f*
  vase outline profiles, 9–11, 9*f*
  Wheatstone cube, 12, 13*f*
Ho, A., 325

Hochberg, J., 71, 185, 187–88, 703
Hogarth, W., 8
Hole, G. J., 67
Hollow mask illusion
  computer modeling, 617
  critique, 151–52, 151*f*
  described, 614–15, 614–15*f*
  drug, alcohol effects, 616
  history of, 614
  infants, perception by, 616
  reclining, 616, 617*f*
  research applications, 617–18
  rotation, 615–17, 615*f*, 617*f*
  strength, factors affecting, 615–16, 615*f*
  variations of, 616–17, 617*f*
Hong, S. W., 379
Hoogstraten, S. van, 7
Horizontal–vertical illusion. *see* vertical-horizontal illusion
Horner, W., 15
Hososhi, 648, 649*f*
Howard, I., 154, 468
Howe, C. Q., 71, 72, 265
Howe, P. D. L., 328, 471
Hoyle, E. M., 69
Hubel, D. H., 202, 207, 556
Hubel, S. H., 99, 104
Hudson, W., 48
Hughes, P., 152, 275–78, 276*f*
Human face, body illusions. *see also* Margaret Thatcher illusion; Presidential illusion; wobbling face illusion
  amodal completion, 255–56, 256*f*
  apparent motion of, 564–66, 565*f*, 567, 567*f*
  clothing, bicolor illusion in, 252, 254*f*
  clothing, Müller-Lyer illusion in, 252, 253*f*
  eye makeup, Delboeuf illusion in, 252, 253*f*
  facial contrast, 620–22, 621–22*f*
  hybrid faces, 775–76*f*, 777
  lightness echo, 255*f*, 257
  magnitude of, 257
  research applications, 622–23
  reversible faces, 633, 635*f*
  shape echo, 254–55, 255*f*, 257
  size contrast, 254–57, 255*f*
Humboldt, Alexander von, 21
Hume, D., 120, 121*f*
Hurlbert, A. C., 383, 385, 394
Hurvich, L. M., 351, 390, 409
Hybrid image illusion
  coarse *vs.* fine-scale information, 772
  described, 763, 764*f*
  facial recognition, 763–64, 765*f*
  local *vs.* global cues, 763, 765*f*
  percept switching, 764*f*, 765
  perceptual hysteresis effect, 765
  psychophysics of, 763
  spatial frequency integration in, 763–65

Ibanez, V., 374
Ibn al-Haytham (Alhazen), 4, 6, 14, 17, 373, 512
Identity hypothesis, 674
Idesawa, M., 510
Illusion of interrupted extent. *see* Oppel-Kundt illusion
Illusion of Sex, 620–22, 621–22*f*
Illusions
  definitions, classification of, xxi, 5, 31–32, 119
  linear perspective rules, 6–7
  nature of, 5–6, 5*f*
  overview, 3–4, 4*f*
  second-order, 404–5*f*, 404–6
  theoretical approaches to, xxi–xxii
Illusions critique
  aftereffects, 150–51
  Ames room (*see* Ames room)
  contrast effects, 149–51
  depth cues, 150, 155n3
  distal stimulus, 145–48, 146–47*f*, 154
  equivalent configurations, 150, 155n3
  facsimile, 148, 154
  filling-in, 151
  hollow masks, 151–52, 151*f*
  information in practice, 148–49, 154, 155n4
  kinetic depth effect, 147
  magnet-like slopes, 147–48, 147*f*, 295–97, 296*f*
  metamerism, 149
  motion parallax, 147
  optic arrays, 144–48
  Ponzo illusion, 153, 154*f*
  proximal stimulus, 145, 155n1
  rate of dilation/time-to-contact, 148
  retinal images, 144–48
  reversed apparent motion, 149–50
  reverspectives (*see* reverspectives)
  sensory-motor adaptation, 154
  simultaneous assimilation, 149–50
  stereoscopes, 146–47, 154
  stimulus, 144
  structure-from-motion transformations, 147
  successive assimilation, 149, 150
  vantage point effects, 147–48, 147*f*
  veridical perception, 144, 146
Illusions of numerosity
  described, 696, 697*f*
  direct comparison, estimation, 697
  estimating process, 698
  impression of power, 698–99
  met (Angetroffene) quality, 698
  parameters, 693–94
  Pegrassi's illusions, 697*f*, 698
  research applications, 696–97
  solitaire illusion, 698, 698*f*
  space judgments, 697
  spatial representations, 698–99
  variants, 693–94
Illusory contours. *see also* subjective contours
  accretions, deletions, 463, 464*f*, 465
  coloration effects, 465, 465*f*
  disk luminance, 465, 466*f*
  occlusion, 463–65, 464–65*f*
  overview, 463
  probabilistic models, 465–66
  stereopsis, 463, 464*f*
Impossible motion illusion, 147–48, 147*f*
Inattentional blindness
  cognition in, 715
  described, 715, 716*f*
  expectations, 716, 717, 717*f*
  familiarity, 716
  individual differences, 716
  noticing rates, 716
  parameters, 715–16
  research applications, 715
  research history, 716–17
  selective looking, 716–17, 717*f*
  similarity, 716
Induced motion, 15–18, 17*f*, 22, 512–15, 513–14*f*, 517, 589. *see also specific illusions by name*
Intersection of constraints model, 82
Invisible saddle. *see* cap or cup illusion
Iris eccentricity hypothesis, 640–44, 641–42*f*, 650n3
Ishii, K., 51
Ishikawa, H., 682*f*
Ito, H., 510, 595n1, 694, 788

Jacobsen, A., 384
Jahoda, G., 44–48, 50, 51
Jakobsson, T., 394
Jameson, D., 351
Jastrow, J., 12, 25
Jastrow illusion, 25, 38–40, 39*f*
Jogging aftereffects, 749–50
Johnston, A., 152
Jones, L., 610
Judd illusion
  length, 206, 206*f*
  Oppel's precursors of, 300–302, 300*f*
  theoretical approaches to, 65*f*, 67, 69

Kamitani, Y., 452
Kamman, R., 195
Kaneko, S., 750
Kang, M.-S., 379
Kanizsa, G., 90–93, 352, 388, 674, 680
Kanizsa square, 91–92, 91*f*, 94, 94*f*, 96, 99, 448, 679, 680*f*, 683–84
Kardos, L., 369
Kastner, K., 304, 306
Katz, D., 370–71
Kayahara, N., 582
Kayahara silhouette illusion
  chirality, 582, 583*f*
  classification, 582, 584
  depth cues, 584
  described, 582, 583*f*
  kinetic depth effect, 582–84, 583*f*
  observer orientation, 582–83, 583*f*
  perspective, 584
  rotation direction, 583–84, 583*f*
  visual hull, 582
Keane, J., 636
Keating, C. F., 47
Keil, M. S., 456
Kellman, P. J., 676, 680, 702
Kelly, F., 92
Kennedy, J. M., 456, 681
Kepler, J., 6, 32
Kersten, D., 383, 385, 394
Kilbride, P. L., 47

Kinetic depth effect
critique, 147
Kayahara silhouette illusion, 582–84, 583*f*
Necker cube, 147
sinusoid illusion, 693, 694
Kingdom, F. A. A., 327
Kingstone, A., 626
Kitaoka, A., 86, 87, 223, 438, 502, 503, 507*f*, 510, 602, 604, 648
Kitayama, S., 51
Klein, B., 194
Knight, E., 440, 772
Koch, C., 383, 721
Koffka, K., 369, 675
Koffka cross, 675
Köhler, W., 102, 778
Komachiya, A., 648
Komatsu, H., 452
Kramer, P., 689, 689*f*
Krekelberg, B., 468
Kriegeskorte, N., 643
Kroes, M. C. W., 467
Krummenacher, J., 685, 686, 686*f*
Kübler, O., 683
Kumar, T., 499
Kundt, A., 27, 305
Kunnapas, T. M., 42
Kuriki, I., 510
Kurtenbach, A., 468

La dame à la licorne, 610, 610*f*
LAMINART theory, 92, 102, 112–13, 113*f*
Landy, M. S., 584, 670
Langlois, M., 397*f*
Lankheet, M. J. M., 467
Lappin, J. S., 488
La Reproduction interdite, 611, 612*f*
La Source, 324–25, 324–25*f*
Lateral inhibition
Chevreul illusion (luminance ramp), 356–60, 356*f*
curved grid non-illusions, 362–64, 363*f*
as mental representation, 127, 128*f*
Necker cube (reversed contrast), 340, 340*f*, 351–52
Laurentini, A., 582
Lawson, R., 610
Lazeyras, F., 374
Leaning tower illusion
acute-angle expansion, 221*f*, 223, 224*f*
compensation for perspective, 222–23, 222*f*
limitations of, 223, 225*f*
overview, 221, 221*f*
perspective view, 221–22
size constancy, 222*f*, 223, 224*f*
unconscious inference, 221–22
Least squares flow estimation, 81–84
Leaver, A. M., 769
Lee, David, 148
Lee, K., 254
Leibowitz, H. W., 46, 47, 49
Leonardo da Vinci, 6, 14, 15, 216, 682
Leonards, U., 374, 456
Leopold, D. A., 655
Levay, S., 749
Leviant, I., 403, 496, 499, 601
Li, C.-Y., 208
Li, Z., 192
Lier, R. van, 293, 438, 452
Limulus crab, 741, 742
Ling, S., 381
Lingelbach, B., 401
Lingelbach, E., 401
Lingelbach's scintillating grid. *see* scintillating grid
Liu, Z., 374
Livingstone, M. S., 99, 104, 207, 468, 471
London, Jack, 136n1
Loomis, R. J., 44
Lorenceau, J., 468, 492, 493, 690
Lotto, R. B., 342
Lu, H., 374
Lu, Z. L., 404, 529, 551–52, 730
Lucas, A., 188
Luccio, R., 696, 698
Lucretius, xx, 16
Luminance and color. *see also* surface lightness; *specific illusions*
afterimages, 323, 323–24*f*
flicker augmented contrast, 325–26, 325*f*
visual pathways, 323–25, 324–25*f*
Luminance contrast polarity
binding effect, 293
boundary conjunction, 292–93, 292*f*
checkerboards, 292, 292*f*
contour-completion phenomenon, 291, 291*f*
contour interpolation, 290–91, 291*f*
distortion, 292*f*, 293
edges, 290–91, 291*f*
good continuity law, 292, 292*f*
logical slant cues, 291–92*f*, 293
mortar line, 291–92, 291*f*
principles, 290, 293
tilts, 291, 291*f*

Macaques, 468, 487
MacDonald, J., 787, 789
Mach, E., 12, 20, 21, 30, 31*f*, 337, 382–84, 394, 404
Mach bands, 404–5*f*, 404–6. *see also* Chevreul illusion (luminance ramp)
Mach card, 12, 13*f*, 394, 394*f*
Mächtigkeitseindruck (impression of power), 698–99
Mack, A., 716
MacKay, D., 571, 738
MacKay rays, 489, 496, 738, 741–42
Macknik, S. L., 649
MacLeod, D. I. A., 422, 746
MacLin, O. H., 655
Maertens, M., 413
MAEs. *see* motion aftereffects (MAEs)
Magic, magicians, 4
Magnet-like slopes, 147–48, 147*f*, 295–97, 296*f*
Malott, M. K., 55
Malott, R. W., 55
Mamassian, P., 584
Maniatis, L. M., 223
Mannoni, L., 12
Margaret Thatcher illusion
autistic children, perception by, 636
configural processing, 635*f*, 636
described, 633, 634*f*
dog faces, 633
face processing, 633–37, 635*f*
faces are unique hypothesis, 633
facial expressions, 635*f*, 636
holistic *vs.* featural encoding, 634–36, 635*f*
infants, perception by, 636
inversion effects, 633–34, 636
parameters, 636
reversible faces, 633, 635*f*
second-order configural cues, 636, 637*f*
variations, 636
Marquise Casati, 648
Marr, D., 79–80, 106, 148, 151, 195
Martinez-Conde, S., 649
Mather, G., 531, 532, 533
Mathison, J., 616
Matisse, H., 102, 103*f*
Maxwell, J. C., xxii
McAdam, M., 584
McCourt, M. E., 327, 328, 329, 391, 415–20, 416*f*, 418*f*, 444
McGeorge, P., 50
McGinitie, W. H., 48
McGurk, H., 48, 787, 789
McGurk effect
auditory cortex modulation, 789–90
facial features studies, 788
infant studies, 788
multisensory integration, 789
native-language studies, 788–89
overview, 787
strength, automaticity, flexibility, 787–89
touch studies, 788
McKelvie, S. J., 636
McPherson, D. A., 42, 43, 50
Melkman, M. J., 391
Mental representations
aftereffects, 127, 129*f*
ambiguity, 127, 131*f*
amodal completion, 127–30, 132*f*
coarse graining in, 126–27, 127*f*
concepts, terminology, 119–22, 120–22*f*, 136nn1–3
deception, 119
delusion, 119
embodiment, 125, 136nn9–10
Gestalt organization, 125, 126*f*, 127–30, 131–34*f*
grouping, 127, 131*f*
hallucinations, 119, 123–24, 136n13
illusiveness, degrees of, 124
inverse optics, 122–25, 124*f*, 136nn4–10
lateral inhibition, 127, 128*f*
local sign, 125
neural substrate, 125
optical factors in, 125–26, 126*f*, 136n12
of physical world, 123–24, 124*f*
physiological bottlenecks, 125
pictorial space, 131–32, 135*f*
psychoanatomy, 132–35, 137n14

scene, 132, 136*f*, 136n13
space-time construction, 131, 134*f*
structural complexity, 127, 129–30*f*
taxonomy, 125, 136n11
Umwelt, 123
user-interfaces, 124
veridicality, 124–25, 132, 136*f*, 136n8, 136n13
visual field effects, 130, 134*f*
visual space effects, 127, 128*f*
Zöllner illusion, 119, 120*f*
Metaxas, D., 617
Metzger, W., 674, 698
Michelson, C., 46
Michotte, A., 674, 703
Micro-saccades, 602, 710, 712
Miellet, S., 764
Mignolla E., 360
Miles, W., 583
Milner, A. D., 76, 187
Mingolla, E., 493
Minnaert, M., 283, 287
Mirror images. *see* Venus effect
Misperceptions of light
attached shadows, 393
Bayesian inference, 393
chromaticity, 393
chromatic shadow illusions, 396–98, 397–98*f*
contrast, 393
Gelb effect, 395, 396*f*
Hering's shadow illusion, 394–95, 395*f*
luminance, 393
Mach card, 394, 394*f*
overview, 393
reflectance, 393
shadow art illusion, 395, 396*f*
unconscious inference, 393
Miyahara, E., 344, 345
Mollon, J. D., 340, 587, 657, 745, 746
Molyneux, W., 6
Mona Lisa gaze effect
described, 639, 640*f*, 645nn1–2
dynamic observer hypothesis, 639–40, 640*f*
egocentric *vs.* egophobic gazes, 639
eyes orientation in, 642*f*, 643, 644
face eccentricity, 642*f*, 643, 644
gaze direction perception, 640–44, 641–42*f*, 646n3
iris eccentricity hypothesis, 640–44, 641–42*f*, 646n3
mechanisms of, 644–45, 645*f*
perceived gaze direction, 643–44
Wollaston's effect, 642–43, 642*f*, 646n3
Monnier, P., 390*f*, 391
Moon (celestial) illusion, xix–xx, 5–6, 5*f*, 308. *see also* new moon illusion
Moon in daylight, 395
Moore, C. M., 687, 688*f*
Morgan, M. J., 67
Morgan, P., 43
Morikawa, K., 252, 254, 256
Motion aftereffects (MAEs)
adaptation, 589–90
adapted motion sensors, 588–89
attentional modulation, 589
binocular rivalry, 722
described, 586, 587*f*
distribution shift model, 586
dynamic noise test patterns, 589
high phi, 592–93*f*, 592–95, 595*f*, 595n1
history, 586
interocular transfer, 587–88
jogging, 749–50
measurement of, 587
monocular/binocular systems interactions, 588
in motion illusions, 7–8, 18–20, 19*f*, 22
MT activity, 589
neural coding, 588
neural substrate, 586, 588*f*
nulling method, 587
phantom, 588–89
ratio model, 586, 588*f*
research applications, 589
in silencing, 579, 579–80*f*
storage, 589
two-stroke apparent motion, 531–33, 532*f*
Motion illusions
aftereffects, 150–51
ambiguous apparent motion, 475–77, 475–77*f*
anomalous, classes of, 501
aperture problem, 81, 81*f*
apparent (*see* apparent motion)
autokinetic effect, 21–22
bicycle spokes illusion, 469, 469*f*
categories of, 489
causal filters, 81, 84–88, 85–87*f*
chopstick illusion, 479–82, 481–82*f*
color from (*see* color from motion)
computational modeling, 79–80
crossover motion, 467, 468*f*
dædaleum, 15
drift movements, 88, 150
energy model, second-order, 523–24, 524*f*
estimation, limitations of, 80–81, 81*f*, 88, 88n1
eye movements in, 13, 15–18, 17*f*, 20, 22, 87–88, 88n2, 150, 482, 482*f*
flash grab illusion, 483–84, 483–84*f*
flow estimation, bias in, 81–84
flying bugs illusion, 478–79, 478–79*f*
footsteps illusion, 469–71, 469–71*f*, 501
furrow illusoion, 472–73, 472–73*f*
Gregory's theory, 80
history of, 7–8, 13
image formation processes, 79
impossible motion, 147–48, 147*f*
induced motion, 15–18, 17*f*, 22, 512–15, 513–14*f*, 517, 589
linear in two directions, 19–20
local *vs.* global motion, 477–78, 477–78*f*
low-level, 467–73
luminance processing in, 87
motion aftereffect (MAE), 7–8, 18–20, 19*f*, 22
objectless, 573–75, 574–75*f*, 575n1
opponency, 514–15, 514*f*
perception, human vision system in, 80
peripheral motion, 472–73, 472–73*f*
phenakistoscope, 15, 16*f*
Plateau spiral, 19, 19*f*
reversed apparent motion, 149–50
reverse phi, 467–69, 468*f*, 522–25, 523–25*f*
scene interpretation processes, 79
sectored disc, 19*f*, 20
sliding phenomena, 501
sliding rings, 482–83, 483*f*
spatiotemporal filtering modeling, 86–87*f*, 86–88
spatiotemporal model processes, 79
spoked wheels, 14–15
in static patterns, 601–5, 601*f*, 603*f*, 605*f*
stationary dot, 17*f*, 18
statistical issues, 81
stimulus arrangements, 17, 17*f*
stroboscopic motion, 13–15, 14*f*, 16*f*, 22
successive exposure, 15
thaumatrope (wonder turner), 14–15, 14*f*
theories of, 79–80
unconscious inference, 17–18
visual persistence, 14–15, 14*f*
visual vertigo, 20–21
waterfall illusion, 18, 19*f*, 20, 31*f*
zig zag, 471–72, 471*f*
Motion-induced blindness
attention, 712
described, 709, 710*f*
eye movements, 710
filling-in, 710*f*, 712
implicit invisibility, 711–12
luminance, contrast, 710, 711*f*
mask, 710, 710*f*, 713
neural mechanisms, 712–13
parameters, 709–10
pattern disappearance, 713
perceptual organization, 710
perceptual suppression, 712–13
research applications, 709
residual subconscious processing, 713
retinal stabilization, 712
spatial layout, 710
target size, position, movement, flicker, 710
variants, 709–10
visual competition, 380–81
visual disappearance, 712
Motion standstill illusion
blue cones, 571
described, 569
equiluminant, isosalient chromatic motion stimuli, 569–71, 570*f*
high spatial frequency gratings, 569, 570*f*
production of, 569, 570*f*
stereomotion standstill, 569, 571, 571*f*
via motion adaptation, 571
Motokawa, K., 400, 401, 403
Motoyoshi, I., 341
Moulden, B. P., 327, 588
Muldrow, E. S., 48
Muldrow, W. F., 48
Müller, Johannes, 24
Müller-Lyer, F. C., 27, 28*f*, 31*f*, 159

Müller-Lyer illusion
in clothing, 252, 253*f*
contrast, assimilation, 65, 65*f*, 150
cross cultural studies, 38–43, 39*f*, 41*f*, 44–46, 49–51
DeLucia/Hochberg configuration, 71, 71*f*
empirical theory of vision, 71–72
history of, xx, xxii, 27, 28*f*
length, 206, 206*f*
maximum-aperture finger configuration, 76, 77
nonhuman species studies, 55, 57–58, 58*f*, 61
numerosity, 696, 697*f*
Oppel's precursors of, 299–300, 299*f*, 301*f*
perspective drawing, 67–69, 68–69*f*
physiological theories, 65
scene statistics, 71–72, 71*f*
size assimilation theory, 67, 67*f*
size scaling, 65*f*, 67–69
transitivity, 64–65, 65*f*
weighted positional averaging, 159–63, 160*f*, 162*f*
Müller-Lyer illusion (3D)
behavior enhancement via, 188
constancy scaling, 185
diagonal edges, 185
fin angle, 188
fin rotation, 185
haptic modality, 186
level of processing, 185–86
pattern, 185, 185–87*f*
perception-action dissociations, 186–87, 188*f*
practice, 188
in real-world settings, 188
stereoscopy, 186
variants of, 188
Mulligan, J. B., 422
Mundy-Castle, A. C., 43
Münsterberg, H., 308, 716
Murakami, I., 86, 509–10, 517
Musatti, C., 352, 388, 597, 598

Nakayama, K., 422, 682, 682*f*, 694
Nam, J.-H., 671
Naor-Raz, G., 501, 505
Navon, D., 102
Necker, L. A., 8, 9, 9*f*, 31*f*, 693, 784
Necker cube
ambiguity, 127, 131*f*, 151, 152
classification, 582, 583*f*
described, 275
history, 9, 12
kinetic depth effect, 147
observer orientation, 582–83, 583*f*
parallelopipeds, 103, 104*f*, 115
reversal, 12, 48
3D interpretations, 103, 104*f*, 115, 275
Necker cube (3D)
described, 784–85, 784–85*f*
multisensory processes, 784–86, 785*f*
observer position, 784*f*
retinal displacements, 784*f*
reversals, 784–85*f*, 784–86
Necker cube (reversed contrast)
binding by synchrony hypothesis, 352–53
configuration spatial articulation, 351, 352*f*
control experiments, 349, 350*f*
lateral inhibition, 340, 340*f*, 351–52
luminance, 350–51, 351–52*f*
neuron response inhibition, 353
perceived color, 350, 351*f*
perceptual belongingness, 349–55, 350*f*, 353–54*f*
perceptual organization, 352
perceptual spatial articulation, 353–54*f*, 353–55
simultaneous contrast, 349, 350*f*, 351–52
Neff, M., 394
Neisser, U., 716–17
Nekes, W., 12
Nelson, G. K., 43
Neon color spreading
boundaries in visual world, 97–100, 97*f*, 99*f*
color assimilation, 390*f*, 392
watercolor illusion, 426–27, 426*f*, 428–29*f*
Nested squares. *see* Vasarely nested squares
Neumann H., 360
Neural blurring, 388–90, 389–90*f*
New moon illusion
angular separation, 286–87, 286*f*, 288n11
cognitive component, 285, 288
distance information, 286*f*, 287
explanation, 287–88, 287*f*
geometry, 285, 288nn9–10
great circle/straight line, 287–88, 287*f*
moon phases, 283–85, 284*f*, 288nn4–8
perceptual component, 285–88, 286–87*f*, 288n11
perspective, 285–88, 286–87*f*, 288n11
plane definition, 283, 284*f*
principles of, 283, 284*f*, 288nn1–3
terminator angle, 285, 288n5
tilt, 284*f*, 285, 288nn4–9
Newport, R., 154
Newsome, W. T., 104, 107, 111
Newton, I., xxii, 14
Ngrams, xx–xxi
Niimi, R., 376
Nili, H., 643
Nishida, S. Y., 452, 522
Nonhuman species studies
contexts, 56, 56*f*
contrast, assimilation, 58, 58*f*, 60, 60*f*, 61
early studies, 54–56, 55*f*
Ebbinghaus–Titchener circles, 55, 58–59, 59–60*f*, 61
ecology/illusion relationships, 61
evolutionary history, 61
height of rectangles, 55
Jastrow figures, 54
Müller-Lyer illusion, 55, 57–58, 58*f*, 61
overview, 54, 55*f*
photographic perspective, 56–57, 57*f*
Ponzo illusion, 56–57, 56–57*f*, 61
vertical–horizontal illusion, 54–55
Zöllner illusion, 55–56, 55*f*, 59–61, 60*f*
Nulling method, 587
Numerosity. *see* illusions of numerosity
Nundy, S., 342

Objectless motion, 573–75, 574–75*f*, 575n1
Oblique-line illusion, 489
Obliques, receding, 234–37, 235–36*f*, 237nn1–2
Occlusion illusion, 269–70*f*, 269–71
O'Connell, D., 582
Öğmen, H, 685, 686, 686*f*
Okumura, K., 256
Oliva, A., 763, 766, 777
Olk, B., 626
Omari, I. M., 48
"On Reflection" exhibition, 610
Ooi, T., 584, 687, 688*f*
Oppel, J. J., xx, 19, 26, 26*f*, 27, 298–303
Oppel-Kundt illusion
attractive forces, 305
centration theory, 305
determining factors, 304–5, 304–5*f*
dividers in, 304–5, 304–5*f*
explanation of, 305–6
history of, 27, 298, 303–4, 304*f*
human face, body, 252
illustrations of, 303–4, 304*f*
luminance in, 305
overview, 303
unconscious inferences, 304*f*, 305–6
variants, 306
Optical flow constraint equation, 82
Optokinetic nystagmus, 105
Orbison illusion, 150
Oriented difference of Gaussians (ODOG) model
grating induction, 416, 417*f*, 419–20
White effect, 328–30, 329*f*
Oruç, İ, 509
Osaka, N., 626
Otero-Millan, J., 411
O'Toole, A. J., 655
Otto, T., 685, 686, 686*f*
Ouchi illusion, 83, 84*f*, 501

Pack, C. C., 107, 493
Paffen, C. L. E., 517, 519, 520
Palmer, E. M., 702
Palmer, S. E., 223, 270, 271
Pantle, A., 685
Papathomas, T. V., 177
Paris, J. A., 14
Parker, A. J., 145
Parker, D. M., 234
Parlangeli, O., 290, 292, 293
Parron, C., 59
Partial modal completion hypothesis, 271
Pasieka, W., 328
Pattern persistence theory, 688
Patterson, Juliette, 397*f*
Pavani, F., 219
Pedersen, D. M., 44
Pedestalled motion paradigm, 573–75, 574–75*f*, 575n1
Pegrassi's illusions, 697*f*, 698
Pelah, A., 749
Penrose Triangle, 584

Pentland, A. P., 369
Perceptual equalization theory, 352
Perceptual error. *see* illusions
Perceptual grouping in space and time theory, 689–90, 689*f*
Perceptual organization
  color from motion, 346
  contrast asynchrony, 769–70, 770*f*
  hidden images, 777–78, 777*f*
  as mental representations, 125, 126*f*, 127–30, 131–34*f*
  motion-induced blindness, 710
  Necker cube (reversed contrast), 352
  subjective contours theories, 683–84
  Ternus effect, 687, 688*f*
Perceptual portraits, 774–78, 775–78*f*
Perceptual theater, xix–xx
Pessoa L., 360, 369
Peterhans, E., 683
Piaget, J., 611, 697
Picciano, L., 685
Pick, H. A., 46
Picture surface illusion
  additive effect, 212
  anamorphics, 212
  ellipses, 210*f*, 211–12
  mechanisms of, 209–10, 210*f*
  motion, 210*f*, 211
  parallelepipeds, 210*f*, 211
  parameters, 210–12, 210*f*
  perspective, 209, 210*f*, 212
  pyramids, 210*f*, 211
  response errors, 210–12, 210*f*
  rotation, 210*f*, 211
  texture, 210*f*, 211, 212
  tilts, 210–11, 210*f*
  triangles, 210*f*, 211
Pierce, A. H., 308
Pigeons, 636
Pikler, J., 685
Pinna, B., 100, 431
Pinna-Brelstaff illusion
  asymmetric luminance profile, 537–38*f*, 541, 541*f*
  categorization of, 489
  children, studies of, 491
  converging-diverging effect, 541, 542*f*
  described, 536, 537–38*f*
  explicit directions role, 544, 544–45*f*
  history of, 501
  illusoriness of effect, 536, 537–38*f*, 546
  implicit diagonals, 537*f*, 538–39, 539*f*, 541–44, 542–43*ff*
  implicit *vs.* explicit diagonals, 537–38*f*, 539–41
  intertwining, spiral effects, 541–44, 543*f*
  line orientation in, 491, 498*f*, 546
  luminance gradients, 543*f*, 545–46, 604
  neural mechanisms, 543*f*, 544–46, 546*f*
  orthogonal biases, 544–46, 546*f*
  peripheral vision, 536–38, 537*f*
  rhombic squares, 538*f*, 541
  rotation, 536, 537–38*f*, 540*f*
Pitchroom illusion
  array orientation, 244–45
  basics of, 241–42, 242*f*
  binocular viewing, 243
  eye movements in, 241
  felt direction of gaze, 248–49
  gravity, 249
  great circle model, 246–48, 247*f*
  hand-to-body distance effect, 245–46, 246*f*
  illumination in, 241
  inducer pitch, 241–42, 242*f*
  linear weighted average model, 243
  line length, orientation, 244, 244*f*
  motion effects, 242
  neuromathematical model, 248, 248*f*
  perceived pitch effects, 242–43, 242*f*
  points arrays, 244
  retinal orientation, 242*f*, 243–44, 244*f*
  size effects, 242, 242*f*
  stimulus amount, 243
  summation, averaging, 244, 245*f*
  visual field complexity, 243
  visually perceived eye level, 242–49, 242*f*, 244*f*, 247*f*, 249n1
Plateau, J., 15, 16*f*, 19, 31*f*, 703
Plug, C., 5, 6
Poggendorff, J. C., 27, 31*f*
Poggendorff illusion
  amodal completion, 673, 674*f*
  carrier patterns, 775, 775*f*
  contrast, assimilation, 65, 65*f*
  cross cultural studies, 40–42, 41*f*, 44–45
  crosstalk theory, 69–70, 70*f*
  history of, 27, 31*f*
  length, 206, 206*f*
  modal completion, 66, 66*f*
  orientation, 207
  perspective drawing, 67–69, 68–69*f*
  physiological theories, 66–67, 66*f*
  scene statistics, 72
  transversal collinearity, 70–71, 70*f*
Pokrzywinski, J., 55
Pollack, R. H., 45, 46, 51
Pollanac, R. B., 50
Ponzo, Mario, 27, 28*f*, 31*f*
Ponzo illusion
  contrast, assimilation, 65, 65*f*
  cross cultural studies, 46–47, 49–50
  depth, 223, 224*f*
  history of, 27, 28*f*, 31*f*
  length, 206, 206*f*
  nonhuman species studies, 56–57, 56–57*f*, 61
  numerosity, 696, 697*f*, 698
  perspective drawing, 67–69, 68–69*f*
  in real-world settings, 188
  retinal images, 153, 154*f*
  size scaling, 65*f*, 67–69, 69*f*
Pool and store model, 264–65
Porac, C., 46, 51
Porterfield, W., 17, 18, 20
Pozzo, A., 7
Prandtl, A., 401
Preattentive vision, 668–69, 669*f*
Presidential illusion
  automatic face classification, 631, 631*f*
  described, 628, 629*f*
  external *vs.* internal features, 628–31, 630*f*
  gender effects, 629, 630*f*
  hairstyles effects, 630
  resolution effects, 629–31, 629*f*, 631*f*
Prismatic shaping, 154
Proffitt, D. R., 187, 192, 352
Prosopagnosia, 657
Psychophysicist's microelectrode. *see* motion aftereffects (MAEs)
Ptolemy, 4–6, 8, 14, 16–17, 275, 512, 649n1
Purkinje, J. E., 15, 18–19, 22–24, 23*f*, 31*f*, 586, 774, 775*f*
Purves, D., 71, 72, 265, 342
Pylyshyn, Z. W., 704

Ramachandran, V. S., 104
Rao, V. M., 104
Ratio model, 586, 588*f*
Ratliff, F., 741, 742
Read, S., 381
Receding objects illusion, 223, 225*f*
Receding obliques, 234–37, 235–36*f*, 237nn1–2
Redundancy-reducing hypothesis, 412*f*, 413
Reeves, A., 431
Reichardt, W., 467, 573, 575, 746
Remijn, G. B., 510
Reuning, H., 42, 43, 45
Reversals
  cross cultural studies, 48
  history, 8, 9, 9*f*, 12
  reversed apparent motion critique, 149–50
Reverse phi
  explicit directions role, 544, 544–45*f*
  first order, 467–69, 468*f*, 523, 524*f*
  Fourier theory, 532–33
  high phi, 592–93*f*, 592–95, 595*f*, 595n1
  history of, 531
  second order, 522–25, 523–25*f*
  single *vs.* two fields, 770–71, 771*f*
  third order, 525
Reverspectives
  concavity, 276*f*, 281
  critique, 152–53, 152*f*, 155n6
  "Daydreaming," 277, 280*f*
  depth misassignment, 279–81, 281*f*
  explanations, 279–81, 281*f*
  "Forced into Reverse Perspective," 277, 278, 280*f*
  forced perspective, 275–78, 277*f*
  history of, 275
  illusory motion, 281, 281*f*
  Necker cube (*see* Necker cube)
  planar perspective, 277, 278
  principles of, 275, 276*f*
  properspectives, 275–77, 277*f*
  "Reverspective Versus Perspective," 277–78, 278*f*
  "Sticking Out Room," 275, 276*f*, 281
  termespheres, 278–79, 280–81*f*
  top-down influences, 281
  "Vanishing Venice," 275, 276*f*
"Reverspective Versus Perspective," 277–78, 278*f*
Révész, G., 54
Revolving wheels illusion. *see* Pinna-Brelstaff illusion

Rhesus monkeys, 479, 636
Ricciardelli, P., 625
Ridge-or-rut illusion, 231, 232*f*
Riley, B., 601
Rittenhouse, D., 8
Ritter, A., 689
Rivers, W. H. R., 38–40, 39*f*, 42–44, 48, 51
Rock, I., 281, 683, 716
Rocking-disk illusion, 770–72, 771*f*
Rogers, B. J., 147, 149, 153, 281, 286–87, 438, 467, 531–33, 752
Rogers, C., 286–87
Rogers, G. T., 490
Roget, P. M., 14–15, 31–32, 31*f*
Rohles, F. H. Jr., 388
Roncato, S., 290, 292, 293
Rosenbach effect, 675
Rosenthaler, L., 683
Ross, H., 5, 6, 369
Rotating snakes illusion, 84–88, 85–87*f*. *see also* Fraser-Wilcox illusion
Rotating tilted lines illusion
  aperture problem, 491
  described, 489–91, 490*f*, 496–99, 497–98*f*
  line orientation in, 491, 493, 493*f*, 496–99, 497–98*f*
  misperception of motion, 489
  motion processing units in, 491, 492*f*
  receptive fields size, 492–93
  retina, circle motion expansion on, 491–92, 492–93*f*
  speed variations, 493
  terminators, 491–93, 492–93*f*
Rouse, H., 636
Rowe, J. B., 643
Rubin, E., 8–11, 9*f*, 427, 430, 694
Rubin, N., 683
Ruiz, O., 401

Sagreiya, H., 471
Samuel, D., 55–56
Sander parallelogram
  cross cultural studies, 41–44, 42*f*, 46, 49–51
  numerosity, 696, 697*f*
Sato, H., 341
Sato, T., 341, 486
Sawides, L., 758–59
Scheiner, Christoph, 6
Schizophrenia, 616
Schlegel, A., 411
Schloss, K. B., 270, 271–72
Schölkopf, B., 288n8
Scholl, B. J., 704
Schor, C. M., 510
Schrauf, M., 362
Schröder's staircase, 12, 13*f*, 582–83, 583*f*
Schuller, S., 83
Schweizer, G., 21
Schyns, P. G., 763, 764, 777
Scintillating grids
  components, 401, 402*f*
  described, 362, 363*f*, 364–66, 501
  diagonal-orthogonal, 403, 403*f*
  discovery of, 400, 400*f*
  dynamic movement, 402–3
  Fourier model, 400*f*, 401, 402
  luminance effects, 402
  red dots effects, 402, 402*f*
  retinal receptive fields, 400–401, 401*f*
  rotation effects, 402–3
  D variant, 401, 402*f*
Scott-Samuel, N. E., 456
Scramble illusion
  blackshot mechanism, 670–71, 671*f*
  boundary contour system, 669, 670*f*, 675n1
  described, 668, 669*f*
  history of, 670
  preattentive vision, 668–69, 669*f*
  probability distribution entropy, 669, 669*f*
  texture metamers, 669*f*, 671
  visual inventory of features theory, 668, 669*f*
Second-order illusions
  luminance modulation function, 404, 404–5*f*
  texture, contrast in, 404–5, 404*f*
Sedgwick, H. A., 239
Segall, M. H., 42–46, 49, 51
Sekuler, R., 501, 505
Selective elongation hypothesis, 271
Selective looking, 716–17, 717*f*
Seno, T., 510, 595n1
Separation of systems, 514–15, 514*f*
Serpell, R., 48
Sex, illusion of. *see* Illusion of Sex
Shadow art illusion, 395, 396*f*, 620–22, 621–22*f*
Shadows
  apparent levitation, 216–17, 216*f*
  correspondence problem, 218
  cue weighting, 215*f*, 216–17
  depth cues, 214–17, 215–18*f*, 216n1
  drop shadows, 216
  Hering's shadow illusion, 394–95, 395*f*
  looming, 218
  motion, 217–18, 217*f*
  overview, 214, 219n1
  perceptual use of, 214–16, 215–16*f*
  visual recognition, 214, 215*f*, 218–19
  visual system processing, 218–19
Shaffer, D. M., 194
Shape echo, 254–55, 255*f*, 257
Shape from smear
  amount of compression, 200–201, 200*f*
  blurring effects, 203, 204*f*
  color, 201, 201*f*
  direction of smearing, 203–4, 204*f*
  image processing, 202
  orientation fields, 204–5
  orientation of compression, 200
  overview, 198, 199*f*
  population response, 202–3, 203*f*
  rendering, 202
  slant effect simulation, 199*f*, 201–2
  smear amount increase, 203, 204*f*
  texture compression, 198–202, 199–201*f*
  texture smearing, 203, 204*f*
  3D appearance, 201, 201*f*
Shape from X computations, 84
Shapiro, A. G., 438, 440, 768, 769, 770, 772
Shapley, R., 413
Shear-Heyman, M., 438
Shepard, R. N., 564
Shepard's box illusion, 234–35, 235*f*, 237n1
Shepard tabletop illusion, 239
Shepherd, J., 50
Shevell, S. K., 390*f*, 391
Shifted chessboard pattern
  Bayesian inference, 311
  dividing lines, 310–11
  history, 308
  kindergarten/café wall version, 308, 309*f*
  line segments sequence, 310, 310*f*
  L-shaped constituents, 310, 310*f*
  mortar lines, 308, 309*f*
  nulling, 310
  top-down processing, 310–11, 310*f*
  visual processing, 308–10, 309*f*
Shimojo, S., 452, 517, 682, 709
Shipley, T. F., 680, 702
Shopland, J. C., 784
Shrinkage illusion, 269–70, 269*f*, 272–73, 272*f*
Silencing
  brief window hypothesis, 578–79, 579–80*f*
  change blindness, 577, 577–78*f*
  dot spacing, brightness, 577–78, 577–78*f*
  in head's-up displays, 580
  motion aftereffects, 579, 579–80*f*
  motion types producing, 578
  perceptual grouping, 579, 579*f*
  retinal receptive fields, 578, 578*f*
Silvar, S. D., 46
Silverman, G. H. H., 694
Simons, D. J., 716*f*
Simultaneous tilt contrast, 150
Singer, B., 342
Singer, W., 352–53
Singh, M., 675
Sinha, P., 463, 624, 625, 630
Sinsteden, W. J., 583
Sinusoid illusion
  ambiguous sensory input, 694
  construction, 692, 693*f*
  kinetic depth perception, 693, 694
  moving, 694
  parameters, 693–94
  perceptions description, 692–93, 694*f*
  perceptual rivalry, 693–95
  variants, 693–94
  visual information processing, 693
Size contrast, assimilation. *see also* contrast, assimilation;*specific illusions by name*
  configural shape illusion, 269, 271–72, 271*f*
  dynamic (*see* dynamic illusory size contrast)
  eating behaviors and, 265
  edge-based, 273
  human face, body illusions, 254–57, 255*f*
  inducer conditions, 273
  models, theories of, 264–65

occlusion illusion, 269–70*f*, 269–71
overview, 262, 269
partial modal completion hypothesis, 271
perception-action model, 265
pool and store model, 264–65
selective elongation hypothesis, 271
shrinkage illusion, 269–70, 269*f*, 272–73, 272*f*
spatial averaging model, 273
spatial proximity in, 265
statistical learning hypothesis, 265
Slant perception
constancy, 191–93, 191*f*, 193*f*
decrease effects, 192
elderly studies, 194
encumberment studies, 194
gravity in, 190, 195
modality effects, 193
perceived gaze declination, 194–95, 195*f*
phenomenology, 190, 190*f*
proprioception, 195–96
sociocognitive factors, 194
surface orientation, 190, 190*f*
surface size effects, 193
theories of, 190–91
viewing direction, 192–93, 193*f*
viewing distance, 191–92, 191*f*, 195
website reporting, 195
Slaven, G. A., 50
Sliding rings illusion, 482–83, 483*f*
Smith, T., 46
Smythies, J. R., 738
Snowden, R. J., 470
Solitaire illusion, 698, 698*f*
Solomon, J. A., 341–42, 446
Spanish castle illusion, 323
Spatial averaging model, 273
Spatiotemporal relatability, 701*f*, 702–4
Spehar, B., 444, 445
Spencer, H., 38
Sperling, G., 275, 341–42, 404, 446, 529, 551–52, 671, 702
Spiegel, H. G., 305
Spillmann, L., 83, 400, 468
Spinning disk illusion, 603*f*, 604
Spinning ellipse speed illusion
aperture problem, 170, 170*f*
aspect ratio/bump rotation, 170*f*, 171, 172*f*
form features motion hypothesis, 172–73, 172*f*
rectangles, 170*f*, 171, 172*f*
signal detection hypothesis, 171, 171*f*
stimulus conditions, 170–71, 170–71*f*
Spitzer, H., 413
Stacey, B., 45–47, 51
Staircase Gelb illusion
articulation, 370–71, 371*f*
codetermination, 369
compression, factors affecting, 369–70, 370*f*
constancy, 367–69
decomposition models, 369
described, 367, 368*f*
gamut compression, 367–70, 370*f*
highest luminace rule, 367, 368*f*
insulation effect, 369–70, 370*f*
intrinsic image models, 369
relative illumination, 369
white border effects, 369–70, 370*f*
Stampfer, S., 15
Stanley, D. A., 683
Static pattern motion illusions
coherent rotation perception, 604
described, 601, 601*f*
displacements, 602, 603*f*
eye movements, 601*f*, 602–3, 603*f*, 605
gratings, phase-modulated, 601*f*, 602–3, 603*f*
luminance gradients, 603*f*, 604
motion energy, 601, 604
response modeling, 604–5, 605*f*
retinal image shifts, 601–2
spatial gradients, 602
temporal modulation, 602
2DMD model, 602–5, 603*f*, 605*f*
Stereokinetic phenomena
contour points, 598, 598*f*
cylinder, 599, 599*f*
described, 597
displacement, 598–99, 598–99*f*
perceptual mechanisms, 597
rotating bar, 598–99, 598–99*f*
rotating ellipse, 597–98, 598*f*
solid ellipsoid, 598, 598*f*
Stereopsis
in color from motion, 346–47
critique, 146–47, 154
illusory contours, 463, 464*f*
motion aftereffects, 588
Müller-Lyer illusion (3D), 186
pitchroom illusion, 243
stereoscopes, 146–47, 154
subjective contours and, 682–83, 682–83*f*
Stereo rotation standstill
capture, 727, 729–30
capture by markers, 728
complex, 730–31
contraction, expansion, 729–30
described, 726, 726–28*f*
flickering texture, 727–28
luminance, 730, 731
motion tracking in, 726–27
parameters, 727, 729*t*
partial capture, 727
physical, 727
position capture, 730
pseudo flicker, 729
scaling, 728–29, 728*f*
spiral motion, 730
static texture, 727, 727*f*
texture motion, 731
variants, 726*f*, 727
Stevens, John, 241
Stewart, V. M., 46, 49
"Sticking Out Room," 275, 276*f*, 281
Stimulus onset asynchrony (SOA), 564–65, 565*f*, 567, 567*f*
Stingray illusion, 178–80, 179–80*f*
Sto, A., 611
Stoner, G. R., 480
Strabismus, 723
Stromeyer, C. F. III, 595n1
Stubbs, D. A., 374
Subjective contours. *see also* boundaries in visual world; filling-in; illusory contours
alignment, 681, 681*f*, 683
binocular vision and, 682–83, 682–83*f*
border ownership, 679–81, 680–81*f*
described, 679, 680*f*
entropy contrast, 681, 682*f*, 683
fill in gaps, 679, 680*f*
fMRI studies, 683
luminosity, 681, 681*f*
perceptual organization theories, 683–84
in pictorial vision, 679–81, 680–81*f*
retinal receptive fields, 683
self-sufficiency, 680, 680*f*
sharp discontinuity, 680–81, 680–81*f*
terminations, 681, 681*f*
T-junctions, 679–80, 680*f*
Subjective visual phenomena history, 22–25, 23*f*, 25*f*
Suga, T., 648
Sugihara, K., 147
Sunaga, S., 510
Sun figure, 456, 457*f*
Surface lightness
boundary effects, 317*f*, 318, 321
chromatic induction, 321, 321*f*
concepts, definitions, 315
decompositions, 315, 321
figure–ground phenomena, 320–21
gradients, 321
Kanizsa figures, 316*f*
layers, scission effects, 316–19*f*, 316–22
occlusion, 316, 316*f*, 320
photometric constraints, 317, 317*f*, 318
sinusoidal gratings, 317, 317*f*
stereoscopic depth, surround colors, 315–17, 316*f*
texture effects, 317–18*f*, 317–21, 320*f*
transparency, 315–16, 316*f*
White's effect, 319, 320*f*
Swaminathan, G., 103
Swimmers illusion, 501
Symmetry, uprightness
altered face, 236–37, 236*f*
bistability, 237
depth, 237
gravity, direction of, 237
mirror symmetry, 235–37, 236*f*, 237n2
parallelograms, 235–37, 236*f*
receding obliques, 234–37, 235–36*f*, 237nn1–2
rectangles, 235–37, 236*f*
retinal projections, 234
rotation, tilt, 236–37, 236*f*
trapezoids, 235–37, 236*f*
veridicality, 234
Symons, L., 626
Szeminska, A., 697
Szily, A. von, 18, 19

Tadin, D., 488
Tahir, H. J., 510
Takahashi, K., 376
Tanaka, J. W., 634
Tanzer, M., 75, 77
Tausch, R., 305

Tcheang, L., 145
Termes, R., 278–79, 280*f*
Termespheres, 278–79, 280–81*f*
Ternus, J., 685
Ternus effect
  accommodation theory, 685
  association fields, 690
  attentional pointer account, 690
  described, 685, 686*f*
  differential activation theory, 688
  element contrast, 687
  factors affecting, 687–88, 687–88*f*
  feature-based bias, 687, 687*f*
  group motion, 685–88, 686*f*, 688–89*f*
  history of, 685
  interstimulus intervals (ISI), 685, 686*f*, 687
  lexical information, 688, 688*f*
  nearest neighbor principle, 688
  nonretinotopic processing, 686
  object identity, 685
  pattern persistence theory, 688
  perceived organization, 687, 688*f*
  perceptual grouping in space and time theory, 689–90, 689*f*
  reference framing, 686
  retinotopic processing, 686
  sensory modality interactions, 686–87
  spatial attention, 688
  theories of, 688–90, 689*f*
  two-motion process distinction, 688
Teufel-Wehrhahn color ring, 452
Texture
  attention-generated apparent motion, 527, 528*f*
  boundaries and, 100–102, 100–103*f*
  in cap or cup illusion, 404–5, 404–5*f*
  compression, 198–202, 199–201*f*
  contrast sensing, 337–42, 338–39*f*, 341*f*
  flickering, 727–28
  grouping, 100–102, 100–103*f*
  metamers, 669*f*, 671
  motion, 731
  picture surface illusion, 210*f*, 211, 212
  in second-order illusions, 404–5, 404*f*
  smearing, 203, 204*f*
  static, 727, 727*f*
  stereo rotation standstill, 727–28, 727*f*, 731
  surface lightness effects, 317–18*f*, 317–21, 320*f*
Thiele, A., 471
Thiéry, A., xxi, 300
Thiéry illusion, 300–302, 300*f*
Thompson, P. G., 471, 778
Thompson, R., 154
Tilts
  Ames room window, 181–83, 182*f*
  line, in Enigma illusion, 496–99, 498*f*
  luminance contrast polarity, 291, 291*f*
  new moon illusion, 284*f*, 285, 288nn4–9
  picture surface illusion, 210–11, 210*f*
  rotating lines (*see* rotating tilted lines illusion)
  simultaneous contrast, 150
  symmetry, uprightness, 236–37, 236*f*
Tingbergen, N., 664
Titchener, E. B., 27, 40
Titian, 609
T-junctions
  amodal completion, 673, 674*f*, 679–80, 680*f*
  color assimilation, 328, 391
  Enigma illusion, 498*f*, 499
  Fraser-Wilcox illusion, 501
  subjective contours, 679–80, 680*f*
  wedding cake illusion, 443–44, 443*f*
  White effect, 328, 391
Todorović, D., 328, 360, 643
Toilet of Venus, 610
Tomimatsu, E., 510
Tootell, R. B. H., 589
Torres Straits study. *see* cross cultural studies
Treisman, A. M., 352
Troje, N. F., 584
Trompe l'oeil, 7
Troncoso, Xoana G., 407, 411
Troscianko, T., 374
Troxler fading, 712
Tse, P., 103
Tseng, C.-H., 275, 730
Tsofe, A., 413
Tsuchiya, N., 721
Tsukasaki, T., 51
Tsuruhara, A., 616
Two-stroke apparent motion
  displacements, 531, 532*f*
  DoG function, 533–34, 534*f*
  Fourier theory, 532–33
  interframe interval (IFI) reversal, 531–33, 532*f*
  luminance effects, 532–34, 532*f*, 534*f*
  motion aftereffects, 531–33, 532*f*
  research applications, 532
  temporal impulse response (TIR), 533, 533*f*
Two visual systems hypothesis, 74–77
Tyler, C. W., 422

Ueda, S., 648
Ueyama, E., 195
Ullman, S., 106

van de Grind, W. A., 589
van der Smagt, M. J., 589
"Vanishing Venice," 275, 276*f*
van Ittersum, K., 265
Van Lier, R., 323, 754
Vantrey, P., 43
Van Tuijl, H. F. J. M., 97, 448
Vardabasso, F., 696
Varin, D., 97, 448
Vasarely, V., 407, 601
Vasarely nested squares
  alternating brightness star, 407–13, 408–12*f*
  BOLD signal, 411–13, 412*f*
  center-surround simulations, 408–9, 411*f*
  corner, curve detection, 408–13, 411–12*f*
  described, 407–8, 408–10*f*
  difference of Gaussians (DOG) filter, 407, 409*f*, 411*f*
  history of, 407
  luminance gradient changes, 373–74
  perceptual quantification, 407–8, 410*f*
  points of maximum curvature discontinuities, 413
  redundancy-reducing hypothesis, 412*f*, 413
  shape, depth perception, 412*f*, 413
Vector decomposition, 514, 514*f*
Vector subtraction, 513–14, 514*f*
Venus effect
  artist's exploitation of, 611, 612*f*
  described, 609–10, 609*f*
  mirrors, projection on, 611–12
  mirrors as familiar objects, 611
  naive physics, 612
  reflections, perception of, 610–11, 610*f*
  size judgments, 612
  variations of, 610, 610*f*
  viewpoint knowledge, 611
Vergeer, M., 323
Verstraten, F. A. J., 589
Vertical-horizontal illusion
  cross cultural studies, 41–47, 41*f*, 50, 51
  length, 206, 206*f*
  nonhuman species studies, 54–55
  numerosity, 696, 697*f*
Vetter, T., 655
Vezzani, S., 273
Vicario, G. B., xxi, 32
Vidyasagar, T. R., 104
Vishwanath, D., 239
Visual-delay-dependent motion illusion, 501
Visual inventory of features theory, 668, 669*f*
Visual system
  angle cells, 103
  blind spot, 94–95, 95*f*
  CLEARS processes, 112–13, 113*f*
  complementary processing streams, 98–99, 98*f*, 104–6, 105*f*
  connectivity development, 139
  disparity gradient cells, 103
  edge-detector cells, 202–3, 203*f*
  end-stopped cells, 683
  feature tracking, 106–7
  gaze-estimation process, 624–26, 626*f*
  information in practice, 148–49, 155n4
  information uncertainty hypothesis, 259–60, 259*f*
  iris eccentricity, 624
  LIP circuit, 107, 108–11*f*, 111
  magnocellular pathway, 665
  motion capture, 107–8, 108–9*f*
  MT activity, 589, 665
  on-cell, off-cell neural arrays, 337–42, 338–39*f*, 341*f*
  parvocellular pathway, 665
  physical properties/perception discrepancies, 139–41, 141–42*f*
  pool and store model, 264–65
  population response, 202–3, 203*f*

processes within, 92–93
processing, standard model, 740–41, 741*f*
recurrent competitive field (RCF), 107, 108*f*
retinal jiggle in, 95–96
retinal painting, 701*f*, 704
retinal receptive fields, 149
S-cone responses, 390*f*, 391
segmentation, integration in, 560–61
sensory qualities, 141
shape/motion integration, 665
statistical learning hypothesis, 265
symmetry-breaking, 99
texture contrast sensing, 337–42, 338–39*f*, 341*f*
T-junctions (*see* T-junctions)
two systems hypothesis, 187–88
visually perceived eye level, 242–49, 242*f*, 244*f*, 247*f*, 249n1
visual stimuli ambiguity, 139, 140*f*
X-junction, 395
Visual vertigo, 20–21
Von Bezold effect, 442, 443*f*
Von Bezold spreading effect. *see* color assimilation
von der Heydt, R., 683
Vurpillot, E., 47–48, 50, 305

Wackermann, J., 304, 305, 306
Wade, N., 144, 775*f*
Wagner, D. A., 49, 50
Wagon-wheel illusion
color, described, 549–51, 550–51*f*
equiluminance effects, 551–53, 552*f*
equiluminance with internal motion, 553–54, 553–54*f*
history of, 548
motion phenomena classes, 551–52, 551*f*
shortest path constraint, 563–64
standard, described, 548–49, 549–50*f*
Walker, J., 287
Wallach, H., 106, 458n3, 512, 582, 597
Wallin, J. E. W., 7, 12
Walls, G., 467
Wansink, B., 265
Warden, C. J., 55
Warner, M., 12
Watanabe, K., 376
Watercolor illusion
asymmetric luminance contrast, 434
boundaries in, 99–100, 99*f*, 434, 434*f*
coloration effect, 426–27, 426*f*, 428–29*f*
coloration without figure-ground, 428–30*f*, 431, 433*f*
discoloration illusion, 433, 433*f*
FACADE model, 435
figure-ground segregation, 426–31, 426*f*, 430–31*f*
figure-ground without coloration, 431–33, 433*f*
lighting illusion, 431–33, 433*f*
neural mechanisms, 434–35
object-hole segregation, 426, 426*f*, 433*f*, 434, 434*f*
visual processing, 435
Watson, A. B., 467
Weak identity hypothesis, 674–75
Webster, M. A., 655, 746
Wedding cake illusion
contrast, assimilation, 442–46, 443*f*
described, 442, 443*f*
luminance constraint, 444–46, 445*f*
perceptual scission, 443–44, 443*f*
research application, 444
T-junctions, 443–44, 443*f*
Wehrhahn, C., 376, 467
Weighted positional averaging
automatic centroid extraction model, 161–63, 162*f*
centroid hypothesis, 161
illusion of extent stimuli, 159–60, 160*f*
illusion strength dependencies, 160–61, 160*f*
terminators, perceptual shifts of, 161
Wells, W. C., 21
Wenceslas illusion, 471, 471*f*
Wertheim, A. H., 517, 519, 520
Wertheimer, M., 102, 352, 429, 531, 563, 573, 697
Wexler, M., 595n1
Wheatstone, C., 12, 13*f*, 15, 24, 28–29, 149, 614
Wheatstone kaleidophone, 14, 14*f*
Wheeler, J., 44
Whistler, R., 633, 635*f*
White, M., 327, 328
White, T., 328
White effect
anchoring model, 328
chromatic, 389*f*, 391
contrast/assimilation, 389*f*, 390, 442–46, 443*f*
cortical filters hypothesis, 327
crossover motion, 467, 468*f*
described, 327, 328*f*
dual mechanism model, 327
layer decomposition, 390
luminance constraint, 444–46, 445*f*
ODOG model, 328–30, 329*f*
pattern-specific inhibition, 327
relative depth relations, 327–28
scission model, 328
spatial filtering model, 444
surface lightness, 319, 320*f*
T-junctions, 328, 391
Wiesel, T. N., 202, 556
Wilcox, K. J., 501
Wilde, J., 46
Willen, J. D., 150
Williams, S. M., 342
Willis, T., 20
Window shade illusion, 770–72, 771*f*
Winters, M., 610
Wobbling face illusion
configural processing, 652–53, 653*f*
described, 648, 649*f*
eye dominance, 650–51, 651*f*
eye movements in, 651–53, 652–53*f*
facial feature duplication, 652–53, 653*f*
facial processing, 649–50, 650*f*
history of, 648
inversion effects, 649–50, 650*f*
plasticity, 653
strength of, 648–51, 651*f*
variations, 651, 651*f*
Wollaston's effect, 642–43, 642*f*, 646n3
Woodworth, R. S., 40, 76
Wortley, W., 42, 43, 45
Wraga, M., 187
Wright, E., 12
Wright, W. D., 391
Wu, B., 584
Wundt, W. M., 27, 31*f*, 300, 302, 305
Wundt illusion
behavior enhancement via, 188
cross-cultural studies, 38–40, 39*f*
history of, 27, 31*f*
Oppel's precursors of, 300–302, 300*f*

Xausa, E., 597

Yagi, A., 486
Yang, E., 341
Yantis, S., 689, 689*f*
Yashima Drive Way, 296, 296*f*
Yazdanbakhsh, A., 92, 328, 491–93
Yin, R. K., 633
Yonas, A., 616
Young, A. W., 635, 636
Young, B. M., 50
Young, T., 6
Yovel, G., 630
Yu, K., 688, 688*f*
Yucht, Y., 413

Zaidi, Q., 328, 415
Zanforlin, M., 597
Zavagno, D., 374, 458n3
Zhou, X., 688, 688*f*
Zig zag illusion, 471–72, 471*f*
Zivotofsky, A. Z., 479
Zöllner, J. K. F., 31*f*
Zöllner illusion
carrier patterns, 775, 775*f*
contrast, assimilation, 65, 65*f*, 150
converging-diverging effect, 541, 542*f*
history of, 27, 27*f*
inhibition, disinhibition, 740–41, 741*f*
length, 206, 206*f*
mental representations, 119, 120*f*
nonhuman species studies, 55–56, 55*f*, 59–61, 60*f*
Oppel's precursors of, 301*f*, 302
orientation, 207
physiological theories, 65–66, 66*f*
Zosh, W., 192
Zusne, L., 38