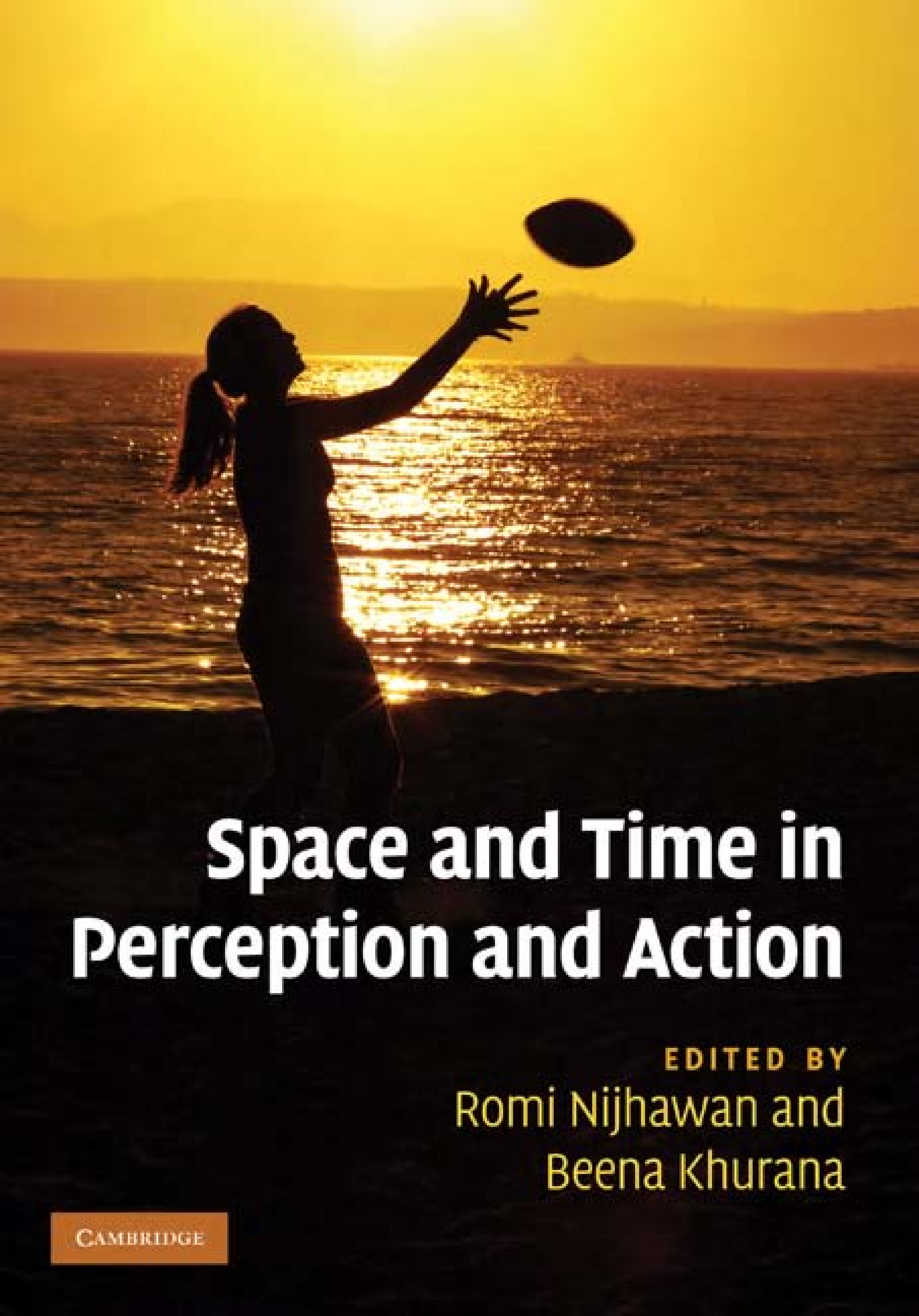


Space and Time in Perception and Action

EDITED BY
Romi Nijhawan and
Beena Khurana

CAMBRIDGE

SPACE AND TIME IN PERCEPTION AND ACTION

What is the instantaneous position of a moving object from the point of view of the observer? How does a tennis player know when and where to place the racket in order to return a 120 mph serve? Does time stop sometimes and go faster at others? Space, time, and motion have played a fundamental role in extending the foundations of nineteenth- and twentieth-century physics. Key breakthroughs resulted from scientists who focused not just on measurements based on rulers and clocks, but also on the role of the observer. Research targeted on the observer's capabilities and limitations raises a promising new approach that is likely to forward our understanding of neuroscience and psychophysics. *Space and Time in Perception and Action* brings together theory and empirical findings from world-class experts and is written for advanced students and neuroscientists with a particular interest in the psychophysics of space, time, and motion.

Romi Nijhawan is a Reader in Psychology at the University of Sussex, UK. In 1994, he introduced the phenomenon and the term "flash-lag effect." He continues to study its implications for the interaction of the animal with the environment.

Beena Khurana is Senior Lecturer in Psychology at the University of Sussex, UK. She is committed to the effective communication of science and has been honored with a Lilly Teaching Fellowship at Cornell University and an Associated Students of CalTech Teaching Award at the California Institute of Technology.

SPACE AND TIME IN PERCEPTION AND ACTION

Edited by

ROMI NIJHAWAN
University of Sussex, UK

BEENA KHURANA
University of Sussex, UK

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,
São Paulo, Delhi, Dubai, Tokyo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521863186

First published 2010

Printed in the United Kingdom at the University Press, Cambridge

A catalog record for this publication is available from the British Library.

Library of Congress Cataloging in Publication data
Space and time in perception and action / edited by Romi Nijhawan, Beena Khurana.
p. ; cm.
Includes bibliographical references and index.
ISBN 978-0-521-86318-6 (hardback)
1. Space perception. 2. Time perception. 3. Space and time. I. Nijhawan, Romi. II. Khurana, Beena.
[DNLM: 1. Visual Perception. 2. Psychophysics. 3. Vision, Ocular. WW 105 S732 2010]
QP491.S64 2010
612′.0145 – dc22 2009036870

ISBN 978-0-521-86318-6 Hardback

Contents

Part I Time–space during action: perisaccadic mislocalization and reaching

Part II Temporal phenomena: perception

Part III Temporal phenomena: binding and asynchrony

Part IV Spatial phenomena: forward shift effects

Part V Space–time and awareness

List of contributors

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

Derek H. Arnold
School of Psychology, The University of Queensland, St. Lucia, Australia

Gisa Aschersleben
Department of Psychology, Saarland University, Saarbrücken, Germany

Holger Awater
Bochum, Germany

Talis Bachmann
Institute of Law, University of Tartu, Tallinn, Estonia

Marcus V. C. Baldo
Institute of Biomedical Sciences, University of São Paulo, São Paulo, Brazil

Harold E. Bedell
College of Optometry and Center for NeuroEngineering and Cognitive Science, University of Houston, Houston, Texas, USA

Eli Brenner
Faculty of Human Movement Sciences, Vrije Universiteit, Amsterdam, The Netherlands

Bruce Bridgeman
Department of Psychology, University of California – Santa Cruz, Santa Cruz, California, USA

David C. Burr
Isitituto di Neuroscienze del CNR, Pisa, Italy

Mark A. Changizi
Department of Cognitive Science, Rensselaer Polytechnic Institute, Troy, New York, USA

Colin W. G. Clifford
School of Psychology, The University of Sydney, Sydney, Australia

David M. Eagleman
Department of Neuroscience, Baylor College of Medicine, Houston, Texas, USA

James T. Enns
Department of Psychology, University of British Columbia, Vancouver, British Columbia, Canada

Wolfram Erlhagen
Departamento de Matemática para C&T, Universidade do Minho, Guimarães, Portugal

Jennifer J. Freyd
Department of Psychology, University of Oregon, Eugene, Oregon, USA

Hiroaki Gomi
NTT Communication Science Labs, Nippon Telegraph and Telecommunication Co., Atsugi, Japan

Patrick Haggard
Institute of Cognitive Neuroscience, University College London, London, UK

Vanessa Harrar
Department of Psychology, York University, Toronto, Ontario, Canada

Laurence Harris
Department of Psychology, York University, Toronto, Ontario, Canada

Hitoshi Honda
Department of Psychology and Center for Transdisciplinary Research, Niigata University, Niigata, Japan

Andrew Hsieh
Pasadena, California, USA

Timothy L. Hubbard
Department of Psychology, Texas Christian University, Fort Worth, Texas, USA

Richard B. Ivry
Department of Psychology, Helen Wills Neuroscience Institute, University of California – Berkeley, Berkeley, California, USA

Philip Jaekl
Universitat Pompeu Fabra, Departament de Tecnologies de la Informació i les Comunicacions, Barcelona, Spain

Dirk Jancke
Institut für Neuroinformatik, Ruhr-University Bochum, Bochum, Germany

Alan Johnston
Department of Psychology, University College London, London, UK

Hulusi Kafaligönül
Vision Center Laboratory, The Salk Institute for Biological Studies, La Jolla, California, USA

Ryota Kanai
Institute of Cognitive Neuroscience, Department of Psychology, University College London, London, UK

Dirk Kerzel
Faculté de Psychologie et des Sciences de l'Éducation, Université de Genève, Genève, Switzerland

Beena Khurana
School of Psychology, University of Sussex, Falmer, UK

Stanley A. Klein
School of Optometry, University of California – Berkeley, Berkeley, California, USA

Christof Koch
California Institute of Technology, Pasadena, California, USA

Agnieszka Kopinska
Department of Psychology, York University, Toronto, Ontario, Canada

Martin J. M. Lankheet
Experimental Zoology Group, Wageningen University, Wageningen, The Netherlands

Markus Lappe
Psychologisches Institut II, Westf. Wilhelms-Universität, Münster, Germany

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

Alejandro Lleras
Department of Psychology, University of Illinois at Urbana-Champaign, Champaign, Illinois, USA

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

Gerrit W. Maus
Center for Mind and Brain, University of California – Davis, Davis, California, USA

Lars Michels
UniversitätsSpital Zürich, Zürich, Switzerland

Geoffrey Miller
Department of Psychology, University of New Mexico, Albuquerque, New Mexico, USA

Cathleen M. Moore
Department of Psychology, University of Iowa, Iowa City, Iowa, USA

M. Concetta Morrone
Isitituto di Neuroscienze del CNR, Pisa, Italy

Ikuya Murakami
Department of Life Sciences, University of Tokyo, Tokyo, Japan

Jochen Müsseler
Psychology Department, RWTH Aachen University, Aachen, Germany

Masayoshi Nagai
Institute for Human Science and Biomedical Engineering, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan

Romi Nijhawan
School of Psychology, University of Sussex, Falmer, UK

Shin'ya Nishida
Human and Information Science Laboratory, NTT Communication Science Laboratories, Atsugi, Japan

Haluk Öğmen
Department of Electrical and Computer Engineering, Center for NeuroEngineering and Cognitive Science, University of Houston, Houston, Texas, USA

Michael Oliver
Vision Science Program, School of Optometry, University of California – Berkeley, Berkeley, California, USA

Saumil S. Patel
Department of Neurobiology and Anatomy, University of Texas Medical School at Houston, Houston, Texas, USA

Gopathy Purushothaman
Department of Cell and Developmental Biology, Vanderbilt Medical Center, Vanderbilt University, Nashville, Tennessee, USA

Leila Reddy
Centre de Recherche Cerveau et Cognition, Faculte de Medecine Rangueil, Toulouse, France

John Ross
School of Psychology, University of Western Australia, Nedlands, Australia

John C. Rothwell
Sobell Department, UCL Institute of Neurology, London, UK

John Schlag
Department of Neurobiology, School of Medicine, University of California – Los Angeles, Los Angeles, California, USA

Madeleine Schlag-Rey
Department of Neurobiology, School of Medicine, University of California – Los Angeles, Los Angeles, California, USA

Shinsuke Shimojo
Computation and Neural Systems, California Institute of Technology, Pasadena, California, USA

Jeroen B. J. Smeets
Faculty of Human Movement Sciences, Vrije Universiteit, Amsterdam, The Netherlands

Mutsumi Suganuma
Global Information and Telecommunication Institute, Waseda University, Tokyo, Japan

Wim A. van de Grind
Helmholtz Institute, Utrecht University, Utrecht, The Netherlands

Rufin VanRullen
Centre de Recherche Cerveau et Cognition, Faculte de Medecine Rangueil, Toulouse, France

Timothy Verstynen
Keck Center for Integrative Neuroscience, Department of Physiology, University of California – San Francisco, San Francisco, California, USA

Katsumi Watanabe
Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo, Japan

David Whitney
Center for Mind and Brain and Department of Psychology, University of California – Davis, Davis, California, USA

Kielan Yarrow
City University, London, UK

Acknowledgments

We thank Rohan Singh Nijhawan for his help with the editing and checking of figures.

This book is based in part on the proceedings of two conferences: Visual Localization in Space–Time held at the University of Sussex, UK (August 2002) and Problems of Space and Time in Perception and Action held at the California Institute of Technology, USA (June 2005). We thank Alisdair Smith (past Vice Chancellor of the University of Sussex); Professor Graham Davey (past Chair, Psychology); the School of Cognitive and Computing Sciences (COGS); EPSRC (EPSRC: GR/S02730/01) for supporting the Sussex conference; and the Association for the Scientific Study of Consciousness (ASSC) for supporting the Caltech conference. Last, we thank the contributors who have made this endeavor worthwhile.

1

Space and time: the fabric of thought and reality

BEENA KHURANA AND ROMI NIJHAWAN

> Space and time are modes by which we think and not the conditions in which we live.
>
> – *Albert Einstein*

Since the beginning of sentience, the fabric of reality has been the subject of intense curiosity, and the twin concepts of *space* and *time* have figured prominently in the thinking of individuals of various intellectual persuasions. Understanding in science has advanced significantly through the postulates that underpin coherence and precision in the representation, and measurement, of space and time. These advances have formed the bedrock of the development of many disciplines. However, until the latter half of the nineteenth century many properties of space and time were assumed and therefore remained unquestioned. For example, the implicit acceptance of concepts such as *absolute space* (a coordinate system at rest, relative to which all inertial frames move at constant velocity) and *absolute time* (a universal time independent of any "clock" or mechanism) made most issues related to space and time impervious to empirical investigation and theoretical debate. This state of affairs was robustly challenged by scientists such as Ernst Mach, who among others imagined observers equipped with measuring devices (rulers and clocks) arriving at concepts at odds with notions of absolute space and absolute time.

Many well-known scientists whose work spanned the latter half of the nineteenth century (Mach included) crossed the disciplinary boundaries of physics, philosophy, and vision science. In Mach's thinking on space and time, the observer's sense perception played a critical role. Mach (1890) wrote: "The facts given by the senses . . . are the starting-point and the goal of all the mental adaptations of the physicist [and] the source of every hypothesis and speculation in science." This statement is reminiscent of another by von Helmholtz (1867): "Apprehension by the senses supplies . . . directly or indirectly, the material of all human knowledge . . . there is little hope that he who does not begin at the beginning of knowledge will ever arrive at its end" (cited in Warren & Warren 1968). It is noteworthy that, although in the new conception of space and time to emerge in the early part of the twentieth century the observer played an integral role, the meaning of the term "observer" remained obscure. Thus, although it was implicit that the observer's nervous system was part of the causal framework, one may ask: Which component(s) of the nervous system

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

are relevant? It is amply apparent that a deeper understanding of space and time, from any point of view, will require a more complete understanding of the observer's nervous system. We suggest that sensory and motor processes in particular not only involve neural representations related to space and time but, more critically, form the basis of the scientist's conception of space and time.

One change apparent in postrelativity thinking is that space and time are no longer thought of as distinct dimensions (Minkowski 1908). We believe it is time for a critical review of separate treatments of space and time in neuroscience and psychophysics. Our unifying efforts are akin to previous efforts by scientists to remove the sharp boundary that is often assumed to exist between perception and action. From a biological standpoint, change and its detection are crucial to the animal's survival. We contend that change, or more generally spatiotemporal events, are the most important stimuli for the nervous system, so it is natural to think of space and time within a unifying perspective.

Change, its detection, and an appropriate response to it are crucial features of all animal behavior. For a single-celled organism, detection of change in the concentration level of some chemical is key to survival. The goal of both internal processes within the animal and its overt behavior in the environment is to maintain homeostasis. For multicelled organisms, change is frequently associated with movement, either because the change itself is due to movement in the environment or because the animal must respond to change with movement. Furthermore, it is well established that change, or any spatiotemporal discontinuity, is a potent stimulus for animal nervous systems. For example, critical information about objects is available at color or luminance edges; stimulus onsets and offsets cause neurons to respond vigorously, whereas static stimuli frequently do not produce a neural response at all, particularly in immature nervous systems; retinal image stabilization (i.e., removal of change) causes visual percepts to disappear rapidly, and so on. Thus, even from a biological standpoint, space and time are naturally connected, and the sharp (intuitive) divide between the two is misleading. It is interesting that a unification of space and time from the point of view of neuroscience and psychophysics seems linked to a unification of perception and action already suggested by a number of notable scientists (e.g., Sperry 1952; Rizzolatti et al. 1997). In conclusion, space and time are connected if one considers moving bodies and clocks from a physical perspective, and when one considers the most significant type of stimulus for biological systems, namely change.

Traditionally, psychologists and neuroscientists treat problems concerning *space* and *time* as more or less two separate and independent problems for investigation. For example, in David Marr's seminal book *Vision,* the spatial aspects of vision were given considerable coverage in stark contrast to the limited analysis accorded to the dimension of time. This is partly because visual pathways are geared to the processing of spatial dimensions. For example, in the domain of space, hyperacuity-level performance (Westheimer 1979) in the two-dimensional plane and in depth appear unrelated to time. Temporal hyperacuity has also been reported (Rose & Heiligenberg 1985). However, such levels of responsiveness require integration over space *and* time. So, the dimension of time is part and parcel of sensory processes even when its role is not apparent or explicitly investigated. Ironically, integration processes that support many types of performance by sensory systems, hyperacuity included, are themselves slow. The large latency of visual processes has been reported time and again (Aho et al. 1993).

On the other hand, of the many articles and books published on the topic of time, few give due consideration to spatial dimensions. One reason for separate empirical investigations of space and time could be that the spatial analysis of events on the one hand, and the timing of events on the other, is carried out by highly distinct mechanisms in the brain. The most natural way to analyze the problem of space is in terms of topographic mapping of the receptor epithelia onto the surfaces of both subcortical and cortical structures, whereas the mechanisms that underlie temporal processing of events may be highly varied. For example, timing mechanisms may be localized in cerebellar processes, motor networks involving the frontal cortex, parietal networks, or some combination thereof. Disparate networks responsible for temporal processing are engaged as a function of the task at hand, with different tasks requiring different networks. These processing differences between space and time may limit an integrated treatment of the two.

Nonetheless, there are important justifications for positing that a unified treatment of space and time is both timely and fruitful. For example, many scientists hold that vision cannot be separated from action, and although vision is often discussed without regard to time, action certainly cannot be similarly divorced from time. There is another aspect to considerations of the dimension of time in vision. Visual perception is not instantaneous; time makes its presence felt in visual processing, particularly because there are significant neural delays at the level of phototransduction and the transmission of receptor signals to the primary visual cortex and beyond. These delays have obvious implications for sensory processes engaged with dynamic visual events and for the motor system's ability to utilize the output of these processes for action. Where change is associated with movement in the visual environment, as is frequently the case, neural delays or issues of time directly impact issues of space. Thus, the relevance of time for vision, already well established for research on auditory processes, is beginning to be appreciated.

A unified treatment of time and space is apparent in the list of chapters. One chief methodology employed by researchers to address problems of space and time involves the study of "illusions," particularly when these illusions are related to or caused by actions. It is well known that animals such as humans are subject to a number of "illusions" related to dimensions of space and time. Action related to perception is thus liable to potential errors unless one takes the strict position that parallel neural streams subserve perception and action. It is doubtful, however, that the two streams are completely independent of each other. In the recent past, researchers have identified and scrutinized several important "illusions." The sixty-four-thousand-dollar questions are: Can the nervous system compensate for these illusions to produce accurate behavioral output? Does the nervous system need to compensate for these illusions, or do these illusions actually aid in the production of adaptive behavior? These questions have become more central to the work on space and time in the last several years. Research conducted in response to such questions is the mainstay of this book.

We bring together theoretical treatments and empirical findings from a number of neuroscientists and psychophysicists with significant experience in the study of space and time. The twin issues of localization in space and time are covered in this book. Two conferences on the topic of Space–Time were key to providing a wellspring of ideas from which this book took shape. The first was titled "Visual Localization in Space–Time" and was held at the University of Sussex (August 2002); the second focused on "Problems of Space and Time in Perception and Action" and was held at the California Institute of Technology (June 2005), as part of the proceedings of the annual conference of the Association for the Scientific Study of Consciousness (ASSC 9). We seek to capitalize on the many fruitful areas of investigation that have emerged in the past several years, and bring together the approaches of scientists who treat time and space as two faces of the same coin (see, e.g., Schlag & Schlag-Rey 2002). The thinking and experiments of researchers working on these topics are presented in a single volume to encourage greater synergism in this exciting field of investigation. This book will achieve its goal if it challenges scientists to bring future questions on space and time under a common umbrella of investigation.

Given that perception is not instantaneous, logic dictates that real-time action must acknowledge and overcome delays inherent in the nervous system. Therefore, we begin with action and the requisite computations of space and time for accuracy in action. Interrogating visual stability in the presence of eye movements has offered insight into the representation of visual space. The late Hitoshi Honda (Chapter 3) deftly analyzes the texture of visual space surrounding a Saccadic eye movement in the presence and absence of visual input. Memory is presented as a cocontributor to vision in maintaining a stable visual world (Lappe, Michels, & Awater, Chapter 4). Using Saccadic eye movements, a case is made for sensorimotor control that requires representations of *both* space and time (Schlag & Schlag-Rey, Chapter 2), whereas a breath of fresh air for psychophysics is presented by the relativistic-like effects of spatial compression and time dilation as a result of shifting gaze (Morrone, Ross, & Burr, Chapter 5). Matin and Li (Chapter 6) make the argument for stability based on a quantitatively precise cancellation function between retinal input and extraretinal position information and the elimination of presaccadic persistence. However, because the underestimation of eye deviation renders compensation via extraretinal signals incomplete, it is provocatively proposed that extraretinal signals are not in the service of compensating prior retinal signals but actually destroying them (Bridgeman, Chapter 7).

What about seeing for reaching? Evidence is sought but none found for object pursuit producing "spatial advanced" representations for overcoming neuromuscular delays (Brenner & Smeets, Chapter 8). At the close of this section the relationship between visual motion and goal-directed reaching is reviewed to conclude that visual motion, although shown to compromise the accuracy of goal-directed reaches, can also contribute to accurate reaching behavior (Whitney, Murakami, & Gomi, Chapter 9).

These initial chapters pivot around representations for action. They are then followed by two sets that focus on temporal and spatial phenomena in perception. We begin with those focused on temporal processing. Going backward in time, a.k.a. temporal antedating, is offered as an account of saccadic chronostasis, or the perceived temporal lengthening of a visual stimulus postsaccade (Yarrow, Haggard, & Rothwell, Chapter 10). Verstynen, Oliver, and Ivry (Chapter 11) measure temporal perception as a function of self-initiated actions to provide compelling evidence for anticipation not only rendering accuracy in action (nothing particularly new there) but critically also affecting our percepts. The result of an investigation into the influence of spatial configurations on perceived durations becomes the basis for spatial priming in temporal kappa effect (Aschersleben & Müsseler, Chapter 12). Then we go on to the perennial problem in vision of establishing coherence out of disjointed sources of information, that is, binding. Temporal binding of visual information has become an intensely researched and fiercely debated enterprise. Clifford (Chapter 13) incisively analyzes both empiricisms and theoretical positions in this burgeoning area of research. Latency differences (Arnold, Chapter 16) and time markers (Nishida & Johnston, Chapter 17) are contrasted with a view that feedback from higher cortical areas to primary visual cortex account for perceived asynchronies. Concordant with this view is the proposal of a high-level brain program for timing based on data for perceived synchronicity of pairs of motion stimuli and pairs of motion and flicker or motion and flashed stimuli (Lankheet & van de Grind, Chapter 18). Eagleman (Chapter 14) attempts to square the issue of processing speed and perceived time by suggesting that the brain computes percepts by waiting for the arrival of the slowest signals. The perception of simultaneity is considered a productive approach to how the brain accurately time stamps events when the process of time stamping itself takes time (Harris, Harrar, Jaekl, & Kopinska, Chapter 15).

Time translates into space. Kerzel (Chapter 19) and Maus (in Maus, Khurana, & Nijhawan, Chapter 27) both bring the classic findings of Fröhlich, Rubin, and Metzger to a wider academic community: The original findings were published in German and remained inaccessible to many interested in the current debates on the spatial and temporal aspects of motion perception. Kerzel attempts to reconcile the Fröhlich effect with the newly discovered onset repulsion effect, whereas Hubbard (Chapter 20) organizes the various theories and models of representational momentum. Based on their assumptions and prowess in terms of accounting for data, Nagai, Suganuma, Nijhawan, Freyd, Miller, and Watanabe (Chapter 21) divvy up representational momentum and the flash-lag effect based on different conceptual influences. The chopsticks illusion offers a window into visual parsing and is used to interrogate whether spatial offsets in the flash-lag effect are computed after motion parsing. Both the chopsticks illusion and reversed phi suggest that the flash-lag effect is a function of motion processing before perception of moving objects (Anstis, Chapter 24).

No understanding of a perceptual phenomenon is complete without a consideration of the role of attention. Baldo and Klein (Chapter 23) carefully scrutinize attention shift delays as modulators and causes of perceived spatial offsets between moving and stationary flashed stimuli. Jancke and Erlhagen (Chapter 25) offer a computation model with biological underpinnings that brings under a shared canopy the Fröhlich, flash-lag, and representational momentum effects. Changizi, Hsieh, Nijhawan, Kanai, and Shimojo (Chapter 26) extend the conceptual canvas further to account for whole classes of geometrical illusions based on the visual system's estimate or "guess" as to the visual syntax of the next moment.

They argue that the visual system's foreknowledge, anchored in exploiting the ecological regularity of forward motion, permits latency correction when forward motion is visually implied. The section closes with an evaluation of various accounts of the flash-lag effect such as differential latency (Kafaligönül, Patel, Öğmen, Bedell, & Purushothaman, Chapter 22), postdiction, and attentional cuing, both theoretically and in light of new data on unpredictable motion (Maus, Khurana, & Nijhawan, Chapter 27).

In the end, one must tackle visual awareness. Enns, Lleras, and Moore (Chapter 28) suggest that perceptual continuity, in the presence of chaotic spatiotemporal inputs, is preserved by operating at the level of objects – object updating. Evidence from visual masking, the flash-lag effect, priming, and perceptual asynchronies is presented in favor of this account. The final two chapters take a stab at the contents of awareness with VanRullen, Reddy, and Koch (Chapter 29) relating the continuous Wagon Wheel illusion to the underlying quasi-periodic brain processes and Bachmann (Chapter 30) offering streamed/continuous stimuli as privy in terms of visual awareness.

We have attempted to be broad and inclusive in our coverage. We hope the diversity of positions adopted in the following chapters, the variety of perceptual phenomena investigated, and the numerous approaches to synthesize first and foremost inform the audience with state of the art in this field. For the future, we look forward to this collection rendering a platform for the problems of space and time in perception and action upon which the next generation of science can build. We tip our hats to Captain Kirk by closing with the thought that space–time is the final frontier in the exploration of our visual world.

References

Aho, A. C., Donner, K., Helenius, S., Larsen, L. O., & Reuter, T. (1993). Visual performance of the toad (Bufo bufo) at low light levels: retinal ganglion cell responses and prey-catching accuracy. *J Comp Physiol [A]* **172**(6): 671–682.

Mach, E. (1890/1897). *Contributions to the Analysis of the Sensations*. Illinois: Open Court.

Minkowski, H. (1908). Space and time. In J. J. C. Smart (ed.), *Problems of Space and Time* (297–312). New York: Macmillan Co.

Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (1997). The space around us. *Science* **277**: 190–191.

Rose, G., & Heiligenberg, W. (1985). Temporal hyperacuity in the electric sense of fish. *Nature* **318**(6042): 178–180.

Schlag, J., & Schlag-Rey, M. (2002). Through the eye slowly: Delays and localization errors in the visual system. *Nat Rev Neurosci* **3**: 191–200.

Sperry, R. W. (1952). Neurology and the mind-brain problem. *American Scientist* **40**: 291–312.

Warren, R. M., & Warren, R. P. (1968). *Helmholtz on Perception: Its Physiology and Development.* New York: John Wiley and Sons.

Westheimer, G. (1979). The spatial sense of the eye. Proctor lecture. *Invest Ophthalmol Vis Sci* **18**(9): 893–912.

Part I

Time–space during action: perisaccadic mislocalization and reaching

2

The internal eye position signal, psychophysics, and neurobiology

JOHN SCHLAG AND MADELEINE SCHLAG-REY

Summary

"Real-time sensorimotor control requires the sampling and manipulation not only of parameters representing space but also of those representing time. In particular, when the system itself has inherent processing delays, it invites a situation in which sampled parameters from a peripheral sensor may no longer be valid at the time they are to be used, due to the change in state that took place during the processing delay" (Dominey et al. 1997). In this chapter, we focus on the situation in which a visual stimulus is flashed near the time of a saccade, and the subject's task is to orient the eyes toward the site where the stimulus has been. To perform this task in complete darkness, the subject's brain has to rely on only two signals: retinal error signal and internal eye position signal (iEPS). This is one of the most interesting situations in which the brain has to compute something in the face of specific physical odds (e.g., very long latencies), and we have some hints on how it proceeds. We analyze the time course of the iEPS – which appears quite distorted – using electrical stimulation of brain structures, instead of natural visual stimuli, to provide the goal to be localized. Different hypotheses are then discussed regarding the possible source and possible neural correlate of the iEPS.

Although vision is usually thought of as a continuous process – continuous in space and time – it is periodically interrupted by rapid eye movements called saccades. These are the movements you make while reading this text. Saccades are necessary because the limited part of the world you see well is the tiny one projecting its image on your fovea. In the retina, only the fovea has a resolution comparable to that of modern digital cameras. Therefore, when a new site of interest appears, you need to reorient your retinal fovea. This may happen several times per second. Each time, even though you don't realize it, your vision becomes transiently vulnerable (much more than during a blink, Deubel et al. 2004). There are several kinds of visual disruptions that may occur at the time of saccades: first, a relative, temporary blindness to changes in the environment (particularly displacements of visual objects) called saccadic suppression (Dodge 1900; Diamond et al. 2000); second, a shift of the apparent location of visual stimuli briefly presented near the time of saccades (Matin & Pearce 1965; Bischof and Kramer 1968; Pola 1976; Mateeff 1978); and third, an apparent bringing closer together of simultaneous visual stimuli, described as space compression (Ross et al. 1997). These three phenomena have about the same time course. As they start appearing for stimuli presented well before saccades (e.g., more than 100 msec) and fading

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

out for stimuli presented well after saccades (e.g., more than 100 msec), clearly their cause cannot be simply mechanical: they cannot be due to the displacement of the eyeball (and the retina with it).

This chapter deals with the perisaccadic mislocalization of brief stimuli that, in darkness, consists in an illusory shift of their position. There are two ways to reveal and measure this shift: one method is perceptual (e.g., locating the test flash with respect to a ruler or a landmark), the other is targeting the visual stimulus (e.g., looking or pointing to the site where it has been). Although initially the results obtained by these two methods were thought to be different (Hallett & Lightstone 1976; Hansen & Skavenski 1985), most recent studies indicate that both methods produce equal mislocalization (Honda 1989, 1990; Dassonville et al. 1992a, 1995; Bockisch & Miller 1999; Boucher et al. 2001). When a shift is observed, its direction and amplitude depend on the timing of the test stimulus. For stimuli presented before the saccade, the shift is in the direction of the saccade. For stimuli presented during the saccade or later, the shift decreases and, for some subjects may even reverse to the opposite direction (e.g., Honda 1990). Complete darkness provides the best condition to demonstrate the shift because, were visual landmarks available, the brain would rely on these cues to recognize spatial relationships (see Bridgeman, this volume), and the phenomenon of compression would also arise (Ross et al. 1997). It is true that mislocalization can still exist in the presence of visual cues and even in the absence of saccades (for instance, when stimuli are moved at saccadic speed in front of an immobile eye, e.g., Ostendorf et al. 2006). But here we are concerned with the saccadic mechanism. Therefore, ideally, the demonstration of perisaccadic shift should be designed with only one test stimulus: preferably a very brief flash and nothing else (no point of fixation and no saccade target). In such an experiment, Dassonville et al. (1995) flashed a point-like stimulus for 2 msec near the time of saccades that subjects had learned to perform upon auditory command. Subjects were instructed to make a first saccade of a given dimension in a specified direction, and then a second saccade to the unmarked site of the flash. In this experiment, targeting was used to measure the mislocalization because any visual cue (e.g., ruler) was to be avoided. A mislocalization of the site of the flash was then observed, which could reach a maximum shift, up to 70% of the first saccade amplitude, when the flash occurred just before saccade onset.

To understand why this mislocalization happens, one should realize that the visual system is slow. It is so slow that it takes more time for visual information to reach the brain than it takes to make a saccade. Therefore, if a flash occurs just before a saccade, the signal that it has occurred is decoded in your brain only after that saccade is completed. How could you ever locate the flash? You could, but only if your brain knows exactly how your gaze was oriented at the instant of the flash. For this purpose, the brain must be able to access – at the proper time – a signal faithfully representing the time course of the saccadic displacement. Is such a signal available? Imagine that you look at a wallpaper that shows a uniformly repetitive pattern of blackberries: how do you know which blackberry you are looking at? Possibly one of the blackberries is perceived with a higher resolution, but this does not help much. You cannot discriminate any blackberry by its visual characteristics because

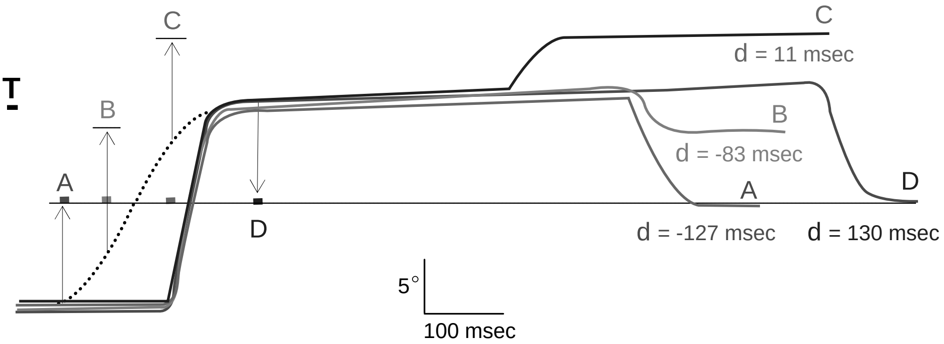


Fig. 2.1 Determination of the iEPS using stimulus mislocalization in complete darkness. Four trials with traces of horizontal eye position recorded by a magnetic scleral coil in a human subject. Two brief stimuli were flashed: T and either A, B, C, or D at indicated times relative to the onset of the saccade to T (= 0 msec). The subject was instructed to make a saccade to T and then to the site where the second flash (A, B, C, or D) was perceived. The delay of the second saccade was not imposed and, therefore, varied in different trials. Black vertical arrows represent retinal vectors. Dotted black curve represents derived iEPS signal. Usually data in such experiments are much noisier than suggested by plotting just a few points as here (see Fig. 2.2 Visual). Other details are described in the text.

all blackberries are the same, and all are at equal distance from each other. Yet apparently you can single out the one you are looking at, as proven if you are allowed to point with your hand (even if you do not see your hand). This seems to imply that there is some sort of a virtual crosshair in your brain that indicates to you where you look. Although this crosshair is invisible, its position is available in terms of its coordinates with respect to the head (and it will be readjusted if you wear prism goggles). This is the internal (or inferred) eye position signal (iEPS). As introduced here, this notion is abstract; we are not referring to any particular biological signal that has yet been identified.

In the absence of other visual cues, the brain can calculate the position of a stimulus briefly presented during a saccade only by adding the iEPS to the retinal position of the stimulus. Absent any internal delay, this addition would be performed in “real time” and the localization of the stimulus could be correct. If it is not correct, we may assume that the iEPS is in error, and from this error we can calculate the iEPS time course. This calculation, explained by the example of Fig. 2.1, is the reverse of the operation specified at the beginning of this paragraph, that is:

$$\text{iEPS} = [\text{retinal position minus saccadic displacement}]$$

Figure 2.1 shows, superimposed, four traces of the horizontal eye position of a subject trying to look successively at the sites of two briefly flashed targets in complete darkness. It is essential that the stimuli be brief because mislocalization diminishes with long stimuli (e.g., $\geq$50 msec; Vliegen et al. 2005). This is not surprising: if stimuli are too long, they start too soon before the saccade (i.e., they are no longer really perisaccadic), or they continue during the saccade and are seen as streaks of light, which helps in their localization. In all four trials the two target locations were the same, but their timing was not. First, target T was always turned on at the start of the trials, whereas the timing of the second target (A, B, C, or D) varied. Thus target A preceded the onset of the saccade to T by 127 msec. Target D followed the offset of the saccade to T by 130 msec. The eye records show that, in both cases, the subject located the second target (A or D) accurately: her gaze movement landed successfully on these sites. The explanation of these successes is straightforward. For target D, the subject had just to register the size of the retinal vector (i.e., distance from stimulus to the fovea, represented in the figure as a black vertical arrow). For finding the position of target A, the subject had two possible solutions. The first one was to take into account the amplitude of the saccade to T and subtract from it the retinal vector to A. The second solution was to estimate the distance between T and A and plan a saccade of that size. Both strategies work. The first one is called *egocentric* because it relies on spatial measurements made with respect to the subject's body (in this case, where the subject was looking). The second one is called *allocentric* (or *exocentric*) because it relies exclusively on relations between stimuli (whatever the body does). Note that if there had not been any target T as in the study of Dassonville et al. (1995), the allocentric solution would not have been possible.

Now consider case B where the target was presented 83 msec before the onset of the saccade to T. Theoretically, the same explanation as proposed for case A should apply here. Yet the subject located this target B much farther away than it really was. Our hypothesis is that the subject was making a considerable error on her estimation of where she was looking when the signal of target B reached her brain. This error is not surprising given the delay of transmitting the position of B to the brain. The error is the difference between the hypothetically inferred eye position and the actual eye position. The landing eye position corresponds to the sum of the retinal vector and the inferred eye position. Note that the allocentric strategy outlined previously would not be able to account for a mislocalization in this case.

The same reasoning can be applied to case C. Here the retinal vector was shorter because the eye had actually already moved when the target was flashed. Again our hypothesis assumes that, when the target was flashed, the subject estimated her eye position to be much closer to the final eye plateau. The four examples illustrated in this figure explain how, by successively testing different stimulus–saccade delays, a curve (dotted line) can be obtained point by point to reconstruct the subject's hypothetically inferred eye position. For this subject, the perisaccadic shift happened to be entirely in the saccade direction. Similar curves are obtained if subjects are asked to report their perception of the stimulus location instead of looking at it (e.g., Matin and Pearce 1965; Mateeff 1978; Honda 1989).

Sperling (1990) approached the notion of iEPS from a slightly different point of view. Because the iEPS corresponds to the apparent displacement of a stimulus that would always be flashed on the fovea during a saccade, he called it the "subjective foveal trajectory." If the objective and subjective foveal trajectories were equal, there should never be any mislocalization. Grüsser et al. (1987) tested this hypothesis directly. They created a foveal afterimage. Afterimages move with the eyes, but at what speed? These authors found that the afterimage moved much more slowly than the eyes. In fact, if subjects made alternate saccades fast enough, the afterimage did not seem to move at all!

As Fig. 2.1 shows, the iEPS curve has two particularities that may seem *a priori* bizarre: first, this curve appears to start before the saccade; second, it is damped.

In fact, the iEPS does not necessarily start growing before the saccade. If the graph of Fig. 2.1 gives this impression, it is because hypothetical events *within the brain* (like the instantaneous value of the iEPS) are plotted in the time frame of events in the *external world* (like the flashes and the saccades). In other words, for each flash A–D the graph shows the value of the iEPS with which the signal from the flash will be combined later in the brain (later than represented in Fig. 2.1). How much later? Theoretically, a time equal to the difference between visual latency and eye position latency to the point in the brain where visual and iEPS signals are combined. The difference between these latencies may be smaller if the iEPS is provided by proprioceptive feedback rather than by a corollary discharge of the oculomotor command. If the apparent timing of the iEPS curve is really a matter of internal delays, it may be possible to change it by changing these delays. Boucher et al. (2001) attempted to lengthen the delay by decreasing the intensity of the test stimuli. Even though the visual latency could be lengthened by as much as 40 msec by this procedure, the results were disappointing. As these authors noted, if the iEPS is severely damped, even a 40-msec latency difference may not noticeably affect the time course of the iEPS. A more radical procedure would be to bypass the afferent pathway completely. Perhaps this could be done by applying an electrical stimulation in the brain that would simulate a retinal position signal? Theoretically, the timing of the iEPS with respect to the saccade should then be much closer.

There is good evidence that, in most oculomotor centers of the brain, such as the superior colliculus (SC), the frontal eye field (FEF), the central thalamus, and some parts of the supplementary eye field (SEF), the effect of electrical stimulation consists of setting a goal to be reached by a saccade. Electrical stimulation acts as if it created a local phosphene (Schlag & Schlag-Rey 1987; Dassonville et al. 1992b) and immediately triggered an eye movement toward that phosphene. Like an afterimage, a phosphene has a fixed retinal position and should move with the eyes. Keeping this in mind, an experiment can be designed, similar to the one illustrated by Fig. 2.1, but in which an electrical stimulation is substituted for the presentation of a flash. Whereas a flash generates a retinal vector, an electrical stimulation generates something equivalent: a saccadic goal at a constant distance in a constant direction with respect to the site where the eyes are pointing at the time of stimulation. Therefore, if that stimulation is applied during an ongoing saccade, it should itself induce a saccade toward a goal referred to the exact position of the eyes when the stimulation was applied. When this experiment is done, one finds that, indeed, there is an electrically evoked saccade following the first one, but it does not go precisely to the predicted site. Here again there is a systematic error. As in the case of natural visual stimuli, this error can be attributed to the fact that the brain is not quite correct in estimating eye

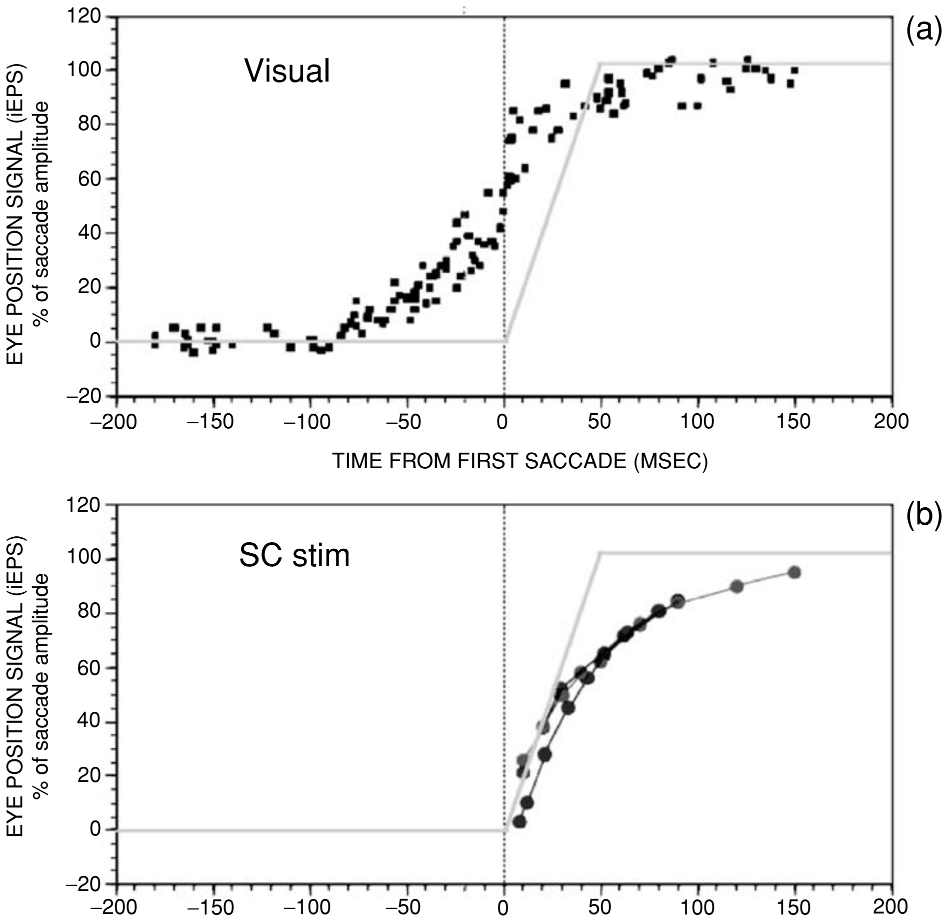


Fig. 2.2 Comparison of iEPS curves obtained in experiments on monkeys using flashed visual targets ((a), Visual) or electrical stimulations of SC ((b), SC stim). Data points were obtained by the procedure illustrated in Fig. 2.1. Abscissa: time relative to onset of first saccade (= 0 msec). First saccade shape drawn in orange. Ordinate: calculated value of iEPS expressed as a percentage of total amplitude of first saccade (100% = 15 degrees). (a) (Visual): data from a single experiment. (b) (SC): as data with electrical stimulation are less noisy than with visual targets, results from three separate experiments could be superimposed. Electrical stimulation applied via implanted tungsten microelectrodes: biphasic (negative first) pulses, 0.2 msec, 200–250 Hz, 140–200 msec train duration, less than 60 μA. The number of stimulus trains applied was limited to avoid lesion.

position during a saccade. Again, the signal that it uses for that purpose, presumably the iEPS, appears to be damped.

In Fig. 2.2, the iEPS curves calculated with stimulation of the SC (b), in monkeys, are compared with the iEPS (in (a)), obtained with natural visual stimuli by the method described by Fig. 2.1. The main behavioral difference between the experiments with visual stimuli and with electrical stimulation is that saccades to visual stimuli occur with longer and more variable latencies (>135 msec in monkey), whereas electrically evoked saccades are triggered by the stimuli (within about 50 msec). Therefore it was not feasible to stimulate before the first (visually guided) saccade in Fig. 2.2(b) because the two saccades (visual and electrically evoked) would overlap and would tend to average their trajectories, as described by Robinson (1972) and Schiller et al. (1979). For this reason, with electrical stimulation only the late part of the iEPS could be tested. Because the subjects of these experiments were monkeys, one does not know what they perceived. Assuming that their perception was consistent with the location of the goal of the evoked saccade, they should have experienced an illusory shift in the direction opposite the saccade (the shift is in this opposite direction when the value of the iEPS is smaller than the corresponding instantaneous value of the actual eye position).

In Fig. 2.2, the curve in (a) (Visual) appears to be sigmoidal as expected if there were random variations. The presence of considerable noise also supports this hypothesis. By comparison, the start of the curves with SC electrical stimulation (b) is more abrupt as if the effect of stimulation had been immediately stronger. However, what seems most important is that the curves in (b) are displaced to the right, as expected if the input visual pathway were bypassed in producing a goal for the eye movement. Results similar to those shown in B have been obtained with electrical stimulation of the FEF (Dassonville et al. 1992a) and SEF (Park et al. 2006), but with slightly longer time constants than for SC (in (b)); further studies can verify this point and possibly explain the difference. These data are preliminary: they are still too imprecise to indicate the exact timing of the iEPS with respect to the saccade. Does it start before or after? This information would be useful to interpret the iEPS as a premotor signal, a corollary discharge, or proprioceptive feedback. One of the difficulties is that we do not know exactly when an electrical stimulation triggers a neuronal response. Indeed, the first few pulses do not seem sufficient because, alone, they cannot produce a saccade. In addition, for more valid comparisons, these experiments need to be performed with standard parameters in the same monkey.

Let us turn our attention to the second strange characteristic of the iEPS curve illustrated in Figs. 2.1 and 2.2: its damping. We will approach this problem from different angles; therefore, the hypotheses considered here are not mutually exclusive.

The internal signal of eye position is usually considered derived from a corollary discharge rather than proprioceptive feedback (e.g., Lewis et al. 1998). Whatever the case, our first hypothesis will be that the signal is filtered. To simulate the course of presaccadic mislocalization, nonlinearities have been introduced in various models to attenuate the slope of saccadic displacements. There may be different explanations for such nonlinearities: for instance, the saccadic feedback coming from a leaky integrator (Dominey et al. 1997), normal variance in visual latency and in saccade characteristics (Boucher et al. 2001), or visual persistence (Pola 2004). As Pola (2004) noted, filtering may be on the visual or eye position path, or both.

Second, the iEPS may not really be stretched in time but appears so because of the procedure used to collect the data points. As we saw (Fig. 2.1), each trial provides only one point to construct an iEPS curve. A large variability due, for instance, to temporal and/or spatial uncertainty (Kanai et al. 2004) would result in an apparent damping of the curve. There is good evidence for temporal and spatial uncertainty at the time of saccade (Morrone et al. 2005), perhaps accentuated by the relative saccadic suppression (Krekelberg et al. 2003).

To this point, the two hypotheses we have considered imply a transformation (possibly even a suspension) of a real signal. Actually, there is a justification for damping: if the brain cannot adequately compensate for internal delays to re-establish the veridical (i.e., external) synchrony of the visual stimulus and the saccade, the best solution is to damp the iEPS. Damping considerably reduces the error of localization (Schlag & Schlag-Rey 1991; Boucher et al. 2001; Pola 2004). This sounds teleological, but it does explain that, whatever causes the damping, the damping solution has reason to survive because it is useful.

Third, there may not be any iEPS available during saccades. Conceivably, the iEPS may be explicit only in static conditions (when the eyes are immobile), momentarily suspended during rapid eye movements, and then restored to its new value when the movements are terminated. During the saccade itself, the firing of eye position neurons might evolve from one level to the next at its own intrinsic rate (Sperling 1990). For instance, something like: the stimulus occurred during the gaze shift from this to that position, and the brain's best guess is that it occurred, say, some time at the beginning of the shift.

Fourth, the iEPS may be a real, meaningful, biological signal – not a distorted copy. For instance, it may be the unaltered representation of the slowly growing *intended* eye position. Thus, it would resemble the Readiness Potential that appears over the scalp of the frontal lobes before voluntary movements. However, it would have one more characteristic property: the amplitude of the signal would correspond to the amplitude of the intended saccade. In many high-level oculomotor centers of the brain, saccades are preceded by long, smoothly progressive increases of firing, sometimes described as preludes. The time course of the iEPS is closer to the time course of such activations than to that of the actual saccadic movement. Can the slow growth of such "prelude-like" signals be revealed by the behavior? Using an ingenious experimental paradigm, Stanford et al. (1990) forced human subjects to make saccades to targets that subjects barely had time to see and locate. Saccades were made with such extremely short latencies that their goal could not be fully determined when they started. With this strategy, Stanford et al. showed that the goal of a saccade is not abruptly, but gradually determined. Interestingly, the time course of this progressive determination is comparable to the time course of curve (b) in Fig. 2.2.

Although there are possible candidates for a neuronal correlate of the iEPS at the time of saccades, there is no good one yet for the perisaccadic shift itself. It would be nice if one could, by unit recording, find activation in the receptive field of a cell (or the receptive fields of an ensemble of cells) that corresponds to the site where the subject illusorily locates the stimulus. One obvious site where such a perisaccadic shift of receptive field could be detected is the lateral intraparietal area in monkey (LIP). In this region, some neurons discharge in response to stimuli that are not yet in their receptive field before a saccade but will be afterward (Duhamel et al. 1992). However, up to now this experiment has only found a correlate of premature *correct* perisaccadic localization (Jeffries et al. 2003; Kusonoki & Goldberg 2003). What is needed is a similar experiment conducted in the particular conditions that, in psychophysical studies, produce perisaccadic *mislocalization* in complete darkness.

Acknowledgment

Supported by USPHS Grant EY-05879.

References

Bischof, N., & Kramer, E. (1968). Untersuchungen und Überlegungen zur Richtungswahrnehmung bei willkürlichen sakkadischen Augenbewegung. *Psychologische Forschung* **32**: 185–218.

Bockisch, C., & Miller, J. M. (1999). Different motor systems use similar damped extraretinal eye position information. *Vision Res* **39**: 1025–1038.

Boucher, L., Groh, J. M., & Hughes, H. C. (2001). Afferent delays and the localization of perisaccadic stimuli. *Vision Res* **41**: 2631–2644.

Dassonville, P., Schlag, J., & Schlag-Rey, M. (1992a). Oculomotor localization relies on a damp representation of saccadic eye displacement in human and nonhuman primates. *Visual Neuroscience* **9**: 261–269.

Dassonville, P., Schlag, J., & Schlag-Rey, M. (1992b). The frontal eye field provides the goal of saccadic eye movement. *Exp Brain Res* **89**: 300–310.

Dassonville, P., Schlag, J., & Schlag-Rey, M. (1995). The use of exocentric and egocentric location cues in saccadic programming. *Vision Res* **35**: 2191–2199.

Deubel, H., Bridgeman, B., & Schneider, W. X. (2004). Different effects of eye blinks and target blanking on saccadic suppression of displacement. *Perception & Psychophysics* **66**: 772–778.

Diamond, M. R., Ross, J., & Morrone, M. C. (2000). Extraretinal control of saccadic suppression. *J Neurosci* **20**: 3449–3455.

Dodge, R. (1900). Visual perception during eye movement. *Physiol Rev* **7**: 454–465.

Dominey, P. F., Schlag, J., Schlag-Rey, M., & Arbib, M. A. (1997). Colliding saccades evoked by frontal eye field stimulaton: artifact or evidence for a compensatory mechanism underlying double-step saccades? *Biol Cybern* **76**: 41–52.

Duhamel, J.-R., Colby, C. L., & Goldberg, M. E. (1992). The updating of the visual representation of visual space in parietal cortex by intended eye movements. *Science* **255**: 90–92.

Grüsser, O.-J., Krizic, A., & Weiss, L.-R. (1987). After-image movement during saccades in the dark. *Vision Res* **27**: 215–226.

Hallett, P. E., & Lightstone, A. D. (1976). Saccadic eye movements due to stimuli triggered during prior saccades. *Vision Res* **16**: 99–106.

Hansen, R. M., & Skavenski, A. A. (1985). Accuracy of spatial localization near the time of saccadic eye movements. *Vision Res* **25**: 1077–1082.

Honda, H. (1989). Perceptual localization of visual stimuli flashed during saccades. *Perception & Psychophysics* **45**: 162–174.

Honda, H. (1990). Eye movements to a visual stimulus flashed before, during, or after a saccade. In M. Jeannerod (ed.), *Attention and Performance XIII: Motor Representation and Control* (567–582). Hillsdale, NJ: Erlbaum.

Jeffries, S. M., Kusunoki, M., Cohen, I. S., & Goldberg, M. E. (2003). Localization errors in double-step saccade task are qualitatively explained by peri-saccadic response patterns in lip. *Soc Neurosci Abstr* **386**: 13.

Kanai, R., Sheth, B. R., & Shimojo, S. (2004). Stopping the motion and sleughting the flash-lag effect: spatial uncertainty is the key to perceptual mislocalization. *Vision Res* **44**: 2605–2619.

Krekelberg, B., Kubischik, M., Hoffman, K.-P., & Bremmer, F. (2003). Neural correlates of visual localization and perisaccadic mislocalization. *Neuron* **37**: 537–545.

Kusonoki, M., & Goldberg, M. E. (2003). The time course of perisaccadic receptive field shifts in the lateral intraparietal area of the monkey. *J Neurophysiol* **89**: 1519–1527.

Lewis, R. F., Gaymard, B. M., & Tamargo, R. J. (1998). Efference copy provides the eye position required for visually guided reaching. *J Neurophysiol* **80**: 1605–1608.

Mateeff, S. (1978). Saccadic eye movements and localization of visual stimuli. *Perception & Psychophysics* **24**: 215–224.

Matin, L., & Pearce, D. G. (1965). Visual perception of direction for stimuli flashed during voluntary saccadic eye movement. *Science* **148**: 1485–1488.

Ostendorf, F., Fischer, C., Gaymard, B., & Ploner, C. J. (2006). Perisaccadic mislocalization without saccadic eye movements. *Neuroscience* **137**: 737–745.

Park, J., Schlag-Rey, M., & Schlag, J. (2006). Frames of reference for saccadic command, tested by saccade collision in supplementary eye field. *J Neurophysiol* **95**: 159–170.

Pola, J. (1976). Voluntary saccades, eye position, and perceptive visual direction. In R. A. Monty & J. H. Senders (eds.), *Eye Movements and Psychological Processes* (245–254). Hillsdale, NJ: Erlbaum.

Pola, J. (2004). Models of the mechanism underlying perceived location of perisaccadic flash. *Vision Res* **44**: 2799–2813.

Robinson, D. A. (1972). Eye movements evoked by collicular stimulation in the alert monkey. *Vision Res* **12**: 1795–1808.

Ross, J., Morrone, M. C., & Burr, D. C. (1997). Apparent position of visual targets during real and simulated saccadic eye movements. *Nature* **386**: 598–601.

Schiller, P. H., True, S. D., & Conway, J. L. (1979). Paired stimulation of the frontal eye fields and the superior colliculus of the rhesus monkey. *Brain Res* **179**: 162–164.

Schlag, J., & Schlag-Rey, M. (1987). Does microstimulation evoke fixed-vector saccades by generating their vector or by specifying their goal? *Exp Brain Res* **68**: 442–444.

Schlag, J., & Schlag-Rey, M. (1991). Spatial programming of eye movements. In J. Paillard (ed.), *Brain and Space* (70–78). Oxford, New York, Tokyo: Oxford University Press.

Sperling, G. (1990). Comparison of perception in the moving and stationary eye. In E. Kowler (ed.), *Eye Movements and Their Role in Visual and Cognitive Processes* (307–351). Amsterdam: Elsevier.

Stanford, T. R., Carney, L. H., & Sparks, D. L. (1990). The amplitude of visually guided saccades is specified gradually in humans. *Soc Neurosci Abstr* **16**: 372–419.

Vliegen, J., van Groetel, T. J., & van Opstal, J. (2005). Gaze orienting in dynamic visual double steps. *J Neurophysiol.* **94**: 4300–4313.

3

Factors influencing perisaccadic visual mislocalization

HITOSHI HONDA

Summary

This chapter is a critical review and discussion of psychophysical studies on perisaccadic visual mislocalization. In particular, it focuses on factors influencing the mislocalization curves. The chapter is organized as follows: first some findings on perisaccadic mislocalization observed in complete darkness are reviewed, followed by empirical and theoretical considerations on eye position signals estimated psychophysically from the mislocalization curves. Next, issues on mislocalization in a lit environment are discussed. Finally, findings on perisaccadic perceptual effects of flickering stimulus are reviewed. Although our understanding of how saccadic eye movements affect visual localization has advanced dramatically in recent years, we probably have only a crude outline of the phenomena and, therefore, further research is needed.

3.1 Visual mislocalization in the dark

3.1.1 Basic findings

In a saccade, the projection of the world sweeps across the retina at high speed. Nevertheless, we usually do not notice this visual motion, and the world continues to appear visually stable. This perceptual phenomenon is called "visual stability."

Although visual stability is preserved during saccades under normal conditions, the perception of the position of objects flashed before, during, or just after a saccade is altered. Examinations into perisaccadic mislocalization were first made in the 1960s by Matin and his colleagues (Matin & Pearce 1965; Matin et al. 1969, 1970). They reported that errors in perceptual localization occurred before the saccade onset and finished slightly after it. Subsequent studies showed essentially the same results (Honda 1989, 1990, 1991; Dassonville et al. 1992, 1995; Sogo & Osaka, 2001). Normally, mislocalization starts about 100–50 msec before the eye begins to move and reaches a maximum at roughly the same time as the saccade onset. Before and during the first half of the saccade, the error is in the direction of the saccade. Later, during the saccade, the error reverses direction, and mislocalization occurs opposite to the saccade direction (Fig. 3.1). Although rebound at the time of saccade offset is sometimes missing (Boucher et al. 2001), this type of mislocalization is usually observed when saccades are executed in the dark.

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

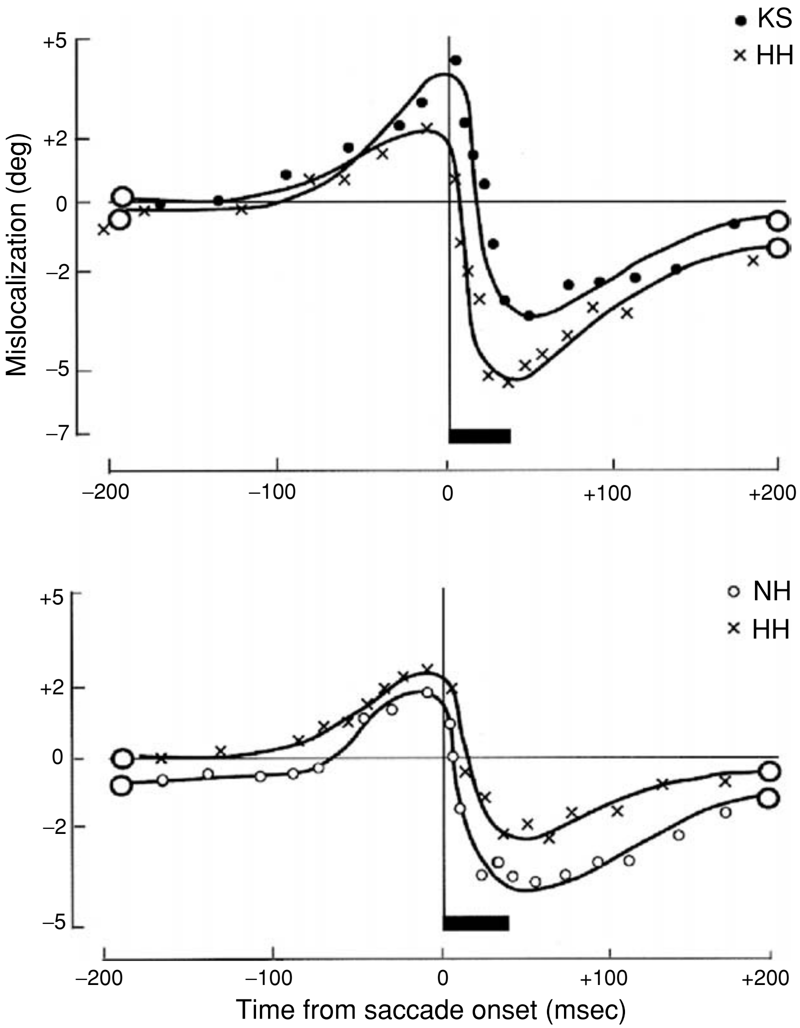


Fig. 3.1 Upper traces show the time course of perceptual mislocalization at the time of horizontal saccades. Two subjects (KS and HH) were required to make an 8 deg saccade in the rightward direction, and a target stimulus for localization was flashed for 2 msec at the position between the original fixation point and the saccade goal. The abscissa indicates the time interval between saccade onset and target presentation. The ordinate indicates the size of mislocalization (deg) of the target. Plus sign in the ordinate shows mislocalization in the saccade direction, and minus sign shows mislocalization in the direction opposite the saccade. Each dot in the figure represents the mean of 10–60 trials. Open circles indicate the results of control trials in which the subjects kept watching the original fixation point (left circle) or the visual cue for eliciting a saccade (right circle). Horizontal bars show the mean duration of the saccades. Lower two traces show the time courses of perceptual mislocalization at the time of vertical saccades. Subjects (NH and HH) were required to make 8 deg saccades in the upward direction. Figure reproduced from Honda (1991) with permission from Elsevier.

A point to note here is that similar mislocalization was observed when a moving light spot was presented at the time of saccades (Kennard et al. 1971). In the experiment in question, a spot of light moved vertically downward on an oscilloscope screen as the eyes moved from left to right and the subjects reported their perceptions by drawing what they saw on the screen. The illusory trajectory of the spot was similar to the mislocalization curve for a light spot flashed at the time of saccades.

Perceptual mislocalization occurs in the same way irrespective of the direction of saccades. Honda (1991) showed that the time course of perceptual mislocalization was the same for vertical and horizontal saccades, despite the difference in the oculomotor system involved in generating these saccades with different directions (Fig. 3.1).

3.1.2 Effects of target position

Honda (1993) examined localization errors observed in five target positions arranged along the path of an 8 deg horizontal saccade and found that the time course of the perisaccadic mislocalization was essentially the same for the five target positions. However, there is no clear consensus on this point. Cho and Lee (2003), for example, presented a target stimulus just after the end of a saccade at one of twenty-five (or eighteen) positions around the saccade goal, and found that the horizontal component of mislocalization increased as the targets were located further into the visual field opposite the saccade. It should be noted here, however, that targets in their experiment were presented for 50 msec, about 20 msec after the end of a saccade, whereas targets in most previous studies were flashed for a few msec at various points in time before, during, or after saccades.

3.1.3 Single saccade versus double saccades

The time course of mislocalization curves discussed so far was obtained for a single saccade. Honda (1997) investigated how the mislocalization curves change when a subject is asked to make a double-step saccade rather than a single saccade. In that study, two types of double-step saccade, the opposite direction and the same direction, were employed. In the former, the second saccade was made toward the original starting position of the first saccade, and the direction of the two saccades was the same in the latter. The subjects were asked to make two saccades successively as fast as possible with a short inter-saccade interval and to report the apparent position of a target flashed at the time of the double-step saccade.

Figure 3.2(a) indicates examples of mislocalization obtained from three subjects for a double-step saccade in the opposite direction. When the inter-saccade interval was short (subjects HH and KN), mislocalization in the direction opposite the first saccade (i.e., downward deflection) continued until the start of the second saccade. On the other hand, when the inter-saccade interval was long (subject TA), mislocalization gradually decreased and disappeared by about 200 msec after the end of the first saccade. Note that the error curve was the same as that previously reported for a single saccade (Fig. 3.1). Thus, mislocalization produced by the first saccade was not influenced by a subsequent second saccade.

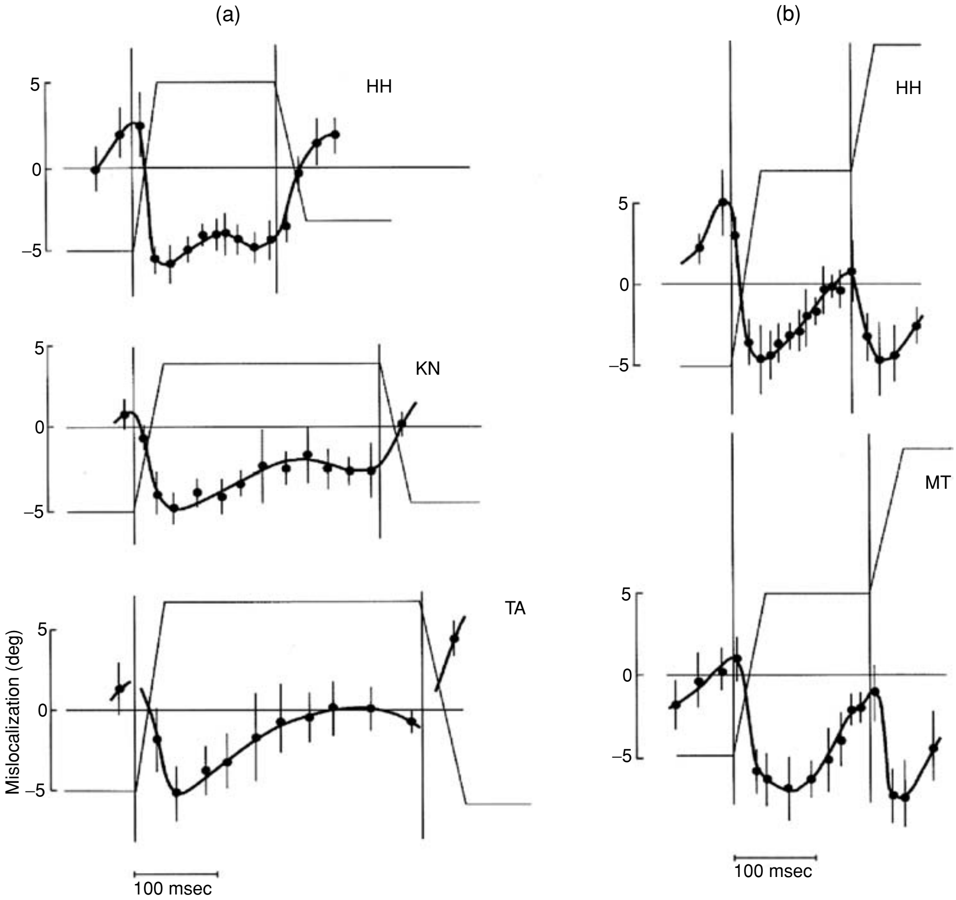


Fig. 3.2 (a) Time course of mislocalization in double-step horizontal saccade trials in which the second saccade was executed in the direction opposite to the first saccade. The distance between the original fixation point and the first saccade goal was 10 deg, and a target stimulus was flashed at the middle between them for 2 msec. The two vertical lines in the figure indicate the onset of the first (left-hand line) and second saccades (right-hand line). The time course of the saccades is schematically shown. The ordinate shows mislocalization (deg). A minus sign in the ordinate indicates mislocalization in the direction opposite the first saccade. Each dot indicates the mean of error of 5–20 trials. (b) Time course of mislocalization in double-step saccade trials in which the second saccade was made in the same direction as the first. The cue for the second saccade was placed 10 deg right of that for the first saccade, and a target stimulus was flashed at the middle of the original fixation point and the cue for the first saccade. Figure reproduced from Honda (1997) with permission from Springer.

Next, we investigated how subjects mislocalize targets when two saccades are made successively in the same direction. Figure 3.2(b) shows two examples of mislocalization curves. The error curves were complex: a large mislocalization in the direction opposite to the first saccade was observed when the target was presented at the end of the first saccade. It rapidly decreased and disappeared immediately before the onset of the second saccade. However, the large error in the direction opposite to the saccade's direction appeared again and reached a maximum at the end of the second saccade.

As discussed later, the localization errors observed for double-step saccades can be used to investigate the temporal interactions of extraretinal eye position signals for two successively executed saccades.

3.1.4 Perceptual versus motor localization

Hallett and Lightstone (1976a,b) reported that the oculomotor responses to a visual target flashed during a prior saccade did not occur on the basis of the retinal position of the target but were directed to the actual target position in space. Hansen and Skavenski (1977, 1985) also indicated that perisaccadic mislocalization was not observed when motor responses, such as hammer blowing, were used instead of perceptual judgments. These findings are intriguing because they are consistent with recent suggestions that there is dissociation between perceptual judgments and motor responses (Milner & Goodale 1995). However, subsequent studies cast doubt on these early findings. Honda (1990) found that eye movements to a visual target flashed at the time of a prior saccade were directed not to the actual target position but to its perceptually misjudged position. Dassonville et al. (1992) also reported that when both human and nonhuman (monkey) subjects were required to make eye movements to the location of perisaccadic visual targets, they showed localization errors similar to those found in perceptual localization studies. Mislocalization was also seen when arm movements were used for localization (Miller 1996; Bockisch & Miller 1999). Although the reason for these discrepancies is not clear, at least for localization errors in the dark, differences in experimental paradigms – especially the availability of possible visual cues introduced from diffuse or reflected light from the apparatus – seem important. For similarities between perceptual and motor localization for conditions with perisaccadic compression, see the chapter by Lappe, Michels, and Awater in this volume.

3.2 Time course of eye position signal

Traditionally, visual stability was explained by a subtraction (or cancellation) mechanism in which any motion of images on the retina is canceled by signals representing the position of the eye (eye position signal; EPS). This idea has often been attributed to von Helmholtz (1866), who thought that the motor command related to desired eye movements is used by the visual system to predict changes of visual input and termed the internal motor command as effort of the will (*Willensanstrengung*). It was later referred to as corollary discharge (Sperry 1950) or efference copy (Von Holst & Mittelstaedt 1950). Because the exact nature of this neural signal is unknown, it is commonly referred to simply as the extraretinal signal.

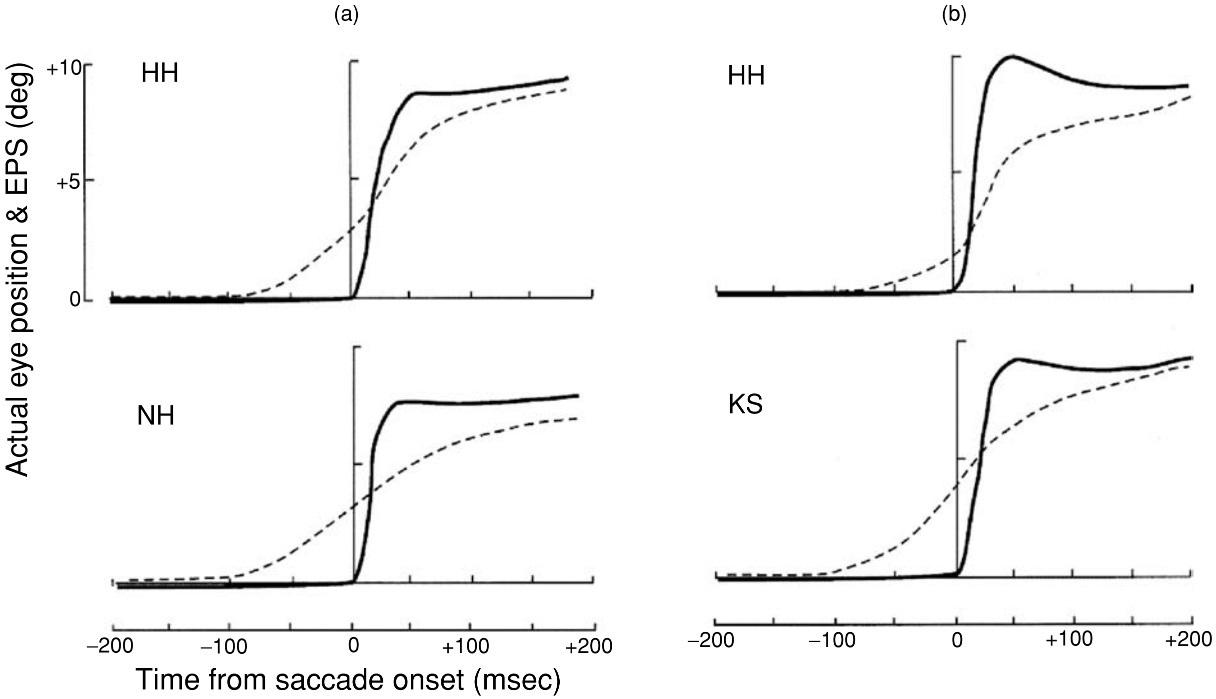


Fig. 3.3 (a) Comparison of the time course of actual saccadic eye movements (solid curves) and that of EPSs (dashed curves) estimated for vertical saccades. (b) The time course of EPSs estimated for horizontal saccades. The ordinate indicates the actual eye position and the estimated magnitudes of EPS (deg). Figure reproduced from Honda (1991) with permission from Elsevier.

3.2.1 EPS for single saccade

The time course of mislocalization shown in Fig. 3.1 is considered evidence that EPS does not represent the physical eye position correctly near the time of saccade execution. According to the subtraction theory, the misperceived location of the flash (P) is determined by summation of its retinal location (R) and EPS. Therefore, EPS can be calculated by $\text{EPS} = \text{P} - \text{R}$ (Mateeff 1978). Figure 3.3 shows an example of the time course of EPS estimated with this equation. As shown in Fig. 3.3, EPS appears approximately 100 msec before the saccade onset but develops so slowly that it cannot catch up with the change of the eye until about 100 msec after the end of the saccade.

Dassonville et al. (1992) reported the same EPS curves for one monkey as well as for human subjects. The authors pointed out that the EPS looks like a “damped” (stretched in time) version of the actual eye position.

3.2.2 EPS for double-step saccade

Honda (1997) estimated the time course of EPS for double-step saccades from the mislocalization curves shown in Fig. 3.2. Figure 3.4(a) shows the time course of the EPS for a double-step saccade estimated on the basis of mislocalization shown by subject HH (Fig. 3.2(a)). The EPS for the second saccade began before the EPS for the first saccade had fully developed. In contrast, when the inter-saccade interval was long (Fig. 3.4(b)), the time course of EPS was the same as that shown for a single saccade (Fig. 3.3). Thus, it is evident that the modification of EPS in double-step saccade trials was seen only in subjects who performed two saccades with a relatively short inter-saccade interval.

Figure 3.4(c) shows the time course of EPS for saccades successively executed in the same direction. As shown in Fig. 3.4(c), EPS began well before the onset of the first saccade and developed slowly. However, it seems that EPS catches up with the eye more quickly than it does in single-saccade trials (Fig. 3.3). Immediately after its arrival at the actual eye position, EPS for the second saccade starts to appear – the EPS for the first saccade rapidly switched over to the EPS for the second saccade without a pause.

Thus, in double-saccade trials, the EPS for each saccade interacted in a specific way, especially when the time interval between the two saccades was short. When the second saccade took place in the direction opposite the first saccade, the EPS for the first saccade was interrupted before completion and switched to the EPS for the second saccade. On the other hand, when the direction of the two saccades was the same, the EPS for the first saccade developed more quickly than when a single saccade was required.

3.2.3 Neurophysiological basis of perisaccadic mislocalization and EPS

Perisaccadic mislocalization occurs well before saccade initiation. This suggests that brain regions active before saccade onset are responsible for generating perisaccadic mislocalization. Neurons in the superior colliculus (SC) (Walker et al. 1995) and the frontal eye

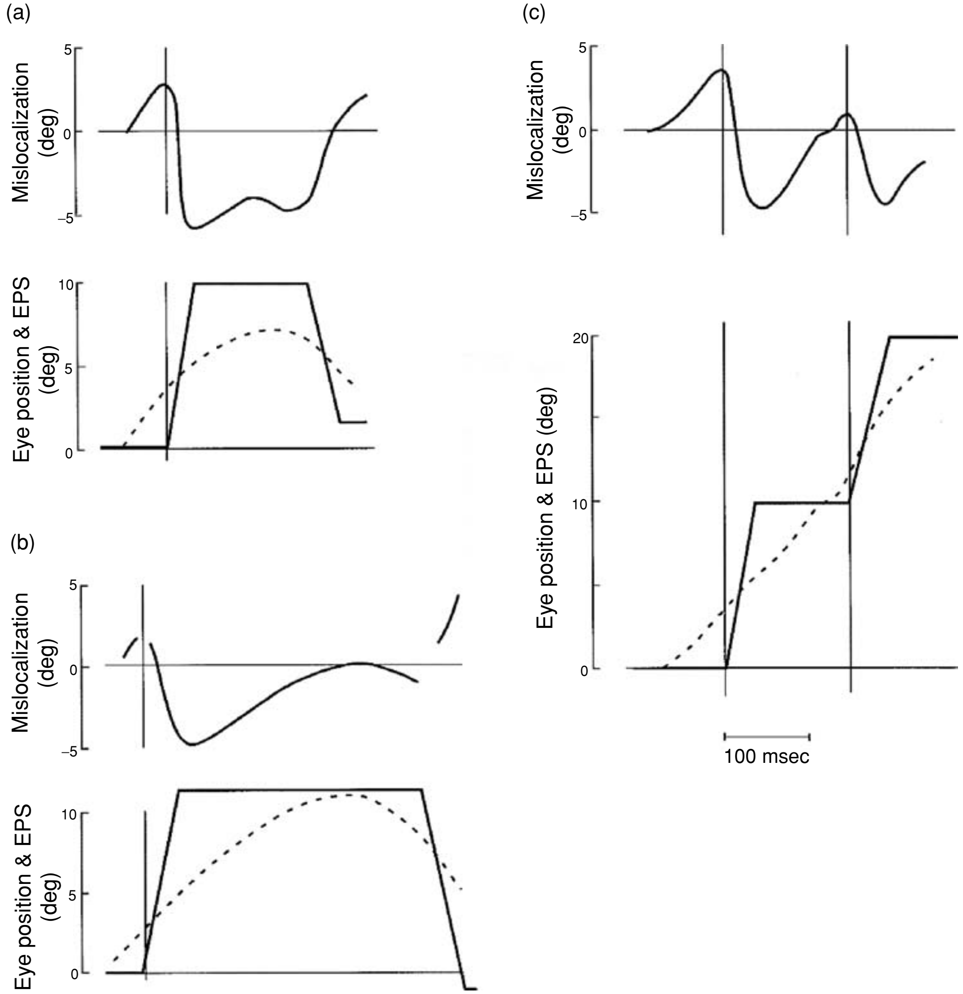


Fig. 3.4 (a) Upper traces show the time course of mislocalization in the double-step saccade trials. Lower two traces show the EPS (dashed curve) and schematically drawn actual eye position (continuous line). Data of subject HH who made two saccades with relatively short inter-saccade intervals. (b) Data of subject TA for whom the inter-saccade interval was relatively long. (c) Upper trace shows the idealized time course of mislocalization drawn on the basis of the data shown in Fig. 3.2(b). Lower two traces indicate the time course of the psychophysically estimated EPS (dashed curve) and schematically drawn actual eye position (continuous line). Figure reproduced from Honda (1997) with permission from Springer.

fields (FEF) (Umeno & Goldberg 1997) show such response properties. Krekelberg et al. (2003) found that perceptual distortion related to saccades can be traced to representation of retinal position in the MT and MST areas. These cells, accurately representing retinal position during fixation, perisaccadically distort the representation of space. In addition, the time course and magnitude of this distortion were similar to the mislocalization found psychophysically in humans. They also suggested an involvement of MT and MST in perisaccadic compression (see 3.3.4).

Receptive fields of neurons in the SC and the FEF shift toward the saccade goal around the time of saccade execution (Walker et al. 1995; Umeno & Goldberg 1997). In the LIP, a shift of receptive fields occurs even before saccade onset (Duhamel et al. 1992). V4 neurons also show a presaccadic shift of receptive fields (Tolias et al. 2001). Thus, these brain areas may be involved in spatial integration across saccadic eye movements.

At present, the neurophysiological mechanism for generating the "damped" EPS is unknown. However, important findings were reported by Sommer and Wurtz (2002). They attempted to identify neurons conveying EPSs (corollary discharge signals for saccades) and hypothesized that the ascending pathway from the SC to the FEF via mediodorsal thalamus (MD) relay neurons carries corollary discharges of saccadic movements. To test this hypothesis, they recorded cell activities in MD relay neurons and found that these cells increase activity before saccade execution. In addition, they examined how inactivation of MD neurons affects a double-step saccade task. In this task, monkeys had to execute successive saccades to two flashed targets, and the ability to make a correct second saccade was thought to rely critically on EPS for the first saccade. The results supported the hypothesis. Inactivation of MD neurons impaired the second saccade consistent with loss of EPS without affecting the first saccade.

It is possible that the mislocalization curve is not caused by the damped EPS but simply reflects errors at the level of neural processing of flashed targets (Schlag & Schlag-Ray 1995). Boucher et al. (2001) investigated the effect of varying neural processing time on localization accuracy for perisaccadic visual targets that differed in luminance. Although systematic errors in localization were observed, the effect of luminance was surprisingly small. Their results indicate that differential neural processing delays, which result from varying stimulus luminance, do not make a large contribution to the pattern of localization errors.

3.3 Visual mislocalization in a lit environment

In the psychophysical experiments discussed so far, saccade-related mislocalization was measured in total darkness. This may be a scientifically useful situation, but certainly not one that we experience in everyday life. Usually exploratory saccades take place in a lit environment.

3.3.1 Basic findings

Bischof and Kramer (1968) and Mateeff (1978) were among the first to measure saccade-related mislocalization in a lit environment. In the Bischof and Kramer study, a target flashed on a horizontal scale with divisions during or after a 16 deg saccade to the right. Subjects were asked to verbally report the scale divisions above which they had seen the target. Clear mislocalization was observed. This depended on the time elapsed between the onset of the saccade and the target presentation. When a target was flashed immediately after the saccade onset, it was mislocalized in the direction of the saccade, whereas when it was presented at the time of saccade offset mislocalization in the direction opposite to the saccade was observed. Mateeff (1978) used an experimental arrangement resembling that used by Bischof and Kramer (1968). In his experiments, subjects performed a 12 deg saccade to the right. The target was presented before, during, or after the saccade, always above the center of the scale. The results were about the same as those reported by Bischof and Kramer (1968). A target presented immediately before or during the saccade was mislocalized in the direction of the saccade. The maximum mislocalization occurred at the saccade onset, and long before and after the saccade the subjects reported the target to be at the veridical position. O'Regan (1984) also conducted experiments in a lit environment. In his experiment, the subjects indicated the position where they saw the flashed target by moving a cursor controlled by a potentiometer knob. He obtained essentially the same results as those reported by Bischof and Kramer (1968) and Mateeff (1978).

In a lit environment, visual cues are abundant, and subjects can use them as visual references in localizing a target. This leads to a reduction in localization errors as compared to those observed in complete darkness. In addition, mislocalization errors are much more influenced by the target position in the visual field (see 3.3.3 and 3.3.5).

3.3.2 Effects of retinal position

Bischof and Kramer (1968) systematically investigated how localization errors are influenced by the position on the retina stimulated. As described in 3.3.1, subjects in their study executed a saccade from left to right, and a target was flashed during or after the saccade. Although there were slight individual differences, the subjects showed similar results. When a target was projected on the hemiretina to which the saccade goal was projected (i.e., when the target was presented on the right side of the fixation position during or after the saccade), subjects tended to mislocalize it at the saccade goal. This mislocalization was observed even when the target was presented immediately after saccade onset. A similar localization error was observed when the target was projected to the fovea. In contrast, when the target was projected on the left side of the fixation position, the target tended to be localized at its veridical position or at the saccade goal. When the target was presented immediately after the saccade onset, it was localized at its actual position; when at the middle or end of the saccade, it tended to be localized at the saccade goal. This rapid displacement of apparent target position from its actual position to the saccade goal occurred immediately after the saccade onset, especially for targets presented near the fovea.

O'Regan (1984) used a computer to track the retina and to project flash targets at predetermined retinal positions, independent of the position the eye reached during a saccade, and confirmed that retinal-location effects do exist. He found that one of his three subjects always located the foveal flash either at the fovea's departure position (i.e., the original fixation position) or at its arrival position (i.e., the saccade goal) and never its in-between veridical position. Interestingly, the foveal flashes that occurred during a saccade were almost always located at the arrival position. Thus, although the direction of gaze changes continuously from one orientation to another during a saccade, the perceived position of the target appears to change discontinuously from the original fixation point to the saccade goal. For nonfoveal flashes, the subject's responses were much closer to veridical. The same differences between foveal and nonfoveal flashes were seen for O'Regan's second subject, but the third subject's data were much noisier. Thus, along with individual differences, the results reported by O'Regan (1984) were somewhat different from those of Bischof and Kramer (1968). In addition, the discontinuous shift of localized position observed in their studies is similar to the discrete jump of a flickering light source in the saccade direction found in Hershberger's (1987) study on phantom array (see 3.4.1).

3.3.3 Effects of flash target position

Honda (1993) investigated mislocalization of a flash target at various positions on a dimly illuminated background. In this study, a map of Japan was drawn on a dark-blue plastic plate with white ink and used as a structured background on which visual stimuli were presented. The luminance levels of the dark-blue area and the white line drawing were about 0.5 and 2 cd/m^2, respectively. Subjects were required to make an 8 deg rightward saccade. A target (40 cd/m^2) for localization was presented at various points in time before, during, or after the saccade. The target position was randomly selected from eleven positions scattered two-dimensionally over the background scene. Subjects indicated the target's apparent position by moving a probe stimulus.

Figure 3.5 shows an example of localization errors in the horizontal direction observed for eleven target positions. The size and direction of the error were dependent on the position at which the targets were presented. When targets were presented at the positions between the fixation point and the saccade goal (i.e., positions [0, +6], [0, +3], [0, 0], [0, –3], and [0, –6] in Fig. 3.5), the error curves were similar to those shown in the dark (Fig. 3.1). However, when targets were presented on the left side of the fixation point (i.e., positions [–8, +8], [–8, 0], and [–8, –3]), localization errors to the saccade direction were prominent. Furthermore, when the targets were presented on the right side of the saccade goal (i.e., positions [+8, +3], [+8, 0], and [+8, –3]), localization errors in the direction opposite the saccade were observed. Finally, mislocalization did not occur in the direction perpendicular to the saccade direction, although a recent study reported mislocalization orthogonal to saccade direction (Kaiser & Lappe 2004).

Subsequent studies have indicated the same results (Honda 1995; Ross et al. 1997). Ross et al. (1997) also found that the perisaccadic mislocalization curve changes with the target position in the visual field. In their experiment, vertical bars were briefly flashed at various positions at various points in time near a 20 deg left-to-right saccade. Subjects reported their location with reference to a ruler. The results were essentially the same as those reported by Honda (1993, 1995). When bars were displayed at 10 deg left of the fixation point or at the position between the fixation point and the saccade goal, they were systematically mislocalized in the direction of the saccade. The shift effects began 50 msec before the

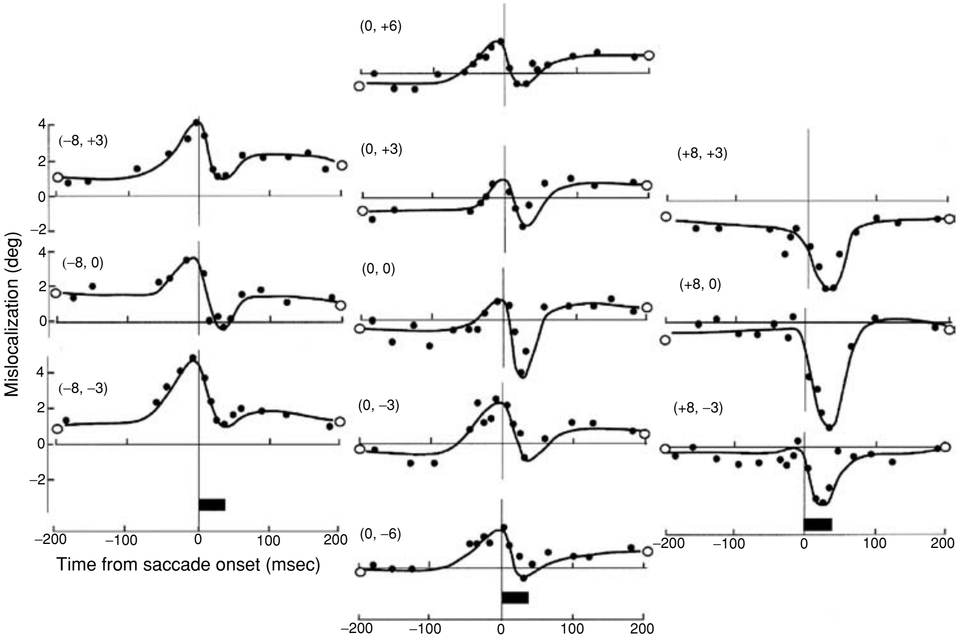


Fig. 3.5 Mislocalization for eleven targets scattered over the dimly illuminated background scene. The figures in the parenthesis indicate the target position. For example, (+8, +3) represents the position 8 deg right and 3 deg above the center of the visual field (0,0). The abscissa indicates the time interval between the saccade onset and the target presentation. The ordinate indicates the size of mislocalization. Plus and minus signs in the ordinate show mislocalization in the saccade direction (toward right) and in the direction opposite to the saccade (toward left), respectively. Each point in the figure represents the average errors calculated for each of the predetermined time intervals of 10 or 30 msec. Open circles indicate the results of control trials in which the subject kept watching the original fixation point (left circle) or the cue for eliciting a saccade (right circle). The duration of the saccades is shown as a horizontal bar. Figure reproduced from Honda (1993) with permission from Elsevier.

saccade, rising to a maximum just before the saccade onset. After the saccade ended, localization was again veridical. When bars were displayed to the right of the saccade goal, the apparent position was displaced against the direction of saccades before eye movement.

3.3.4 Compression of visual space

As mentioned in 3.1.1, Kennard et al. (1971) found that when the eye moved in the horizontal direction across the visual field as a light spot moved downward, a zigzag figure was seen. They furthermore found that shifting the vertically moving light spot to the right or left altered the perceived motion of the light spot. Based on these findings, they predicted that, at the time of saccade onset, objects in the visual field seem not only to be displaced in the saccade direction, but also to be compressed.

Ross et al. (1997) experimentally demonstrated that there is saccade-contingent compression of visual space. In one of their experiments, subjects reported the apparent separation of two half-bars displayed collinearly at different times, one 75 msec after the other. The most interesting finding was obtained when the latter was flashed just before saccade onset. When the bars were presented near the original fixation position, the latter bar was mislocalized with respect to the earlier in the direction of the saccade. On the other hand, when the bars were presented at the position beyond the saccade goal, the effect inverted, showing a shift against the direction of the saccade (see also Brenner et al. 2005). For a stronger test of compression, they displayed up to four bars within a 20 deg region centered on the saccade goal and asked subjects to report how many they saw. Well before or after the saccade, subjects seldom erred, but near the saccade onset, they reported seeing only one bar, no matter how many were displayed. In an additional experiment, they displayed several natural scenes at various intervals relative to a saccade. When they were displayed just before the saccade, all subjects saw the scenes compressed. A subsequent study by Lappe et al. (2000) showed that saccade-contingent compression occurs only when visual references are available immediately after, rather than before or during, the saccade onset. They argued that their findings indicate that the visual processes of transsaccadic localization use mainly postsaccadic visual information.

If there is saccade-contingent compression of visual space, objects located at the saccade goal are expected to appear compressed. Matsumiya and Uchikawa (2001) showed this is not the case. Their subjects were asked to judge the apparent width of a rectangle (a single element) or four bars (four elements) presented just before saccades. The results were that the apparent width of the four-bar stimulus was compressed just before saccades, but the width of the rectangle stimulus was not compressed. These findings indicate that the shape of a single object is not distorted at the saccade goal during presaccadic compression of visual space.

3.3.5 The role of visual reference

In a lit environment, visual cues are abundant, and subjects can use them as visual references in localizing a target flashed at the time of saccades. In other words, they can localize it exocentrically (or allocentrically) with respect to the location of other stimuli in the visual field. This leads to a reduction of localization errors as compared with those in complete darkness in which localization is made mainly on the basis of inaccurate damped EPS. Dassonville et al. (1995) asked their subjects to localize perisaccadic stimuli in the presence or absence of a visual stimulus that could provide exocentric location information. Saccadic localization was more accurate in the presence of exocentric stimuli. Dassonville et al. interpreted their finding as showing that localization is based on a combination of exocentric and EPS based egocentric cues.

Honda (1999) examined how changing the structure of the background visual scene modifies the perisaccadic mislocalization. When targets were presented in complete darkness, a large mislocalization was observed. The pattern of localization errors was similar to that shown in Fig. 3.1. On the other hand, when targets were presented on a dimly illuminated background or within a bright rectangular visual frame drawn against a dark background, localization errors were similar to ones shown in Fig. 3.5. In these conditions, furthermore, the size of mislocalization was reduced in comparison with that shown in the dark. In addition, localization errors in a lit environment consisted of nonuniform mislocalization across the visual field, that is, mislocalization that depends on the spatial position of the target in space.

Awater and Lappe (2004) investigated the role of visual cues for eliciting saccades in producing perisaccadic mislocalization. For this purpose, they used a paradigm known as the antisaccade task (Hallett 1978). Subjects were asked to make a saccade to a visual cue stimulus (prosaccade) or to the mirrored position of the visual cue (antisaccade). Target stimuli were presented in the dark or with a visual reference (a horizontal ruler). The results were clear. The pattern of mislocalization was similar for pro- and antisaccades. Furthermore, localization errors followed actual eye movement and not the visual cue that elicited the saccade. From these findings and previous findings (Lappe et al. 2000), Awater and Lappe concluded that the initial visual cue is not important for perisaccadic localization errors. In addition, they argued that their findings disagree with recent theories on transsaccadic stability in which saccade targets play a crucial role in the perception of space around saccadic eye movements (Deubel et al. 1998; Currie et al. 2000).

Burr et al. (2001) investigated localization errors for verbal (perceptual) reporting and pointing (motor) localization tasks. Under normal lighting conditions, stimuli briefly flashed just before saccade onset were mislocalized in pointing as well as in verbal reporting. When subjects pointed with the eye closed, however, they accurately localized the correct target position. Their study thus presents good evidence for separate systems, one for perception localization and one for motor localization. In addition, mislocalization in the verbal reporting task was observed even when all physical references, including the visual scale, were removed just after target presentation. This suggests that the visual scale recalled from memory served as some form of reference.

3.4 Perceived locations of successively flashed perisaccadic stimuli

3.4.1 Basic findings: A phantom array

As described previously, a perisaccadically flashed target is mislocalized. Based on this localization error, many researchers believe that EPS changes slowly compared with the physical eye position when a saccade takes place (the damped-EPS model). The damped-EPS model is widely accepted, but Hershberger (1987) has argued against this model. His subjects were asked what was perceived when they performed a saccade across a point light source flickering on and off at 120 Hz. Most of his subjects reported that the light source initially jumped in the direction of the saccade, then displaced in the direction opposite to the saccade, and finally stopped near its physical position (Fig. 3.6). Hershberger (1987)

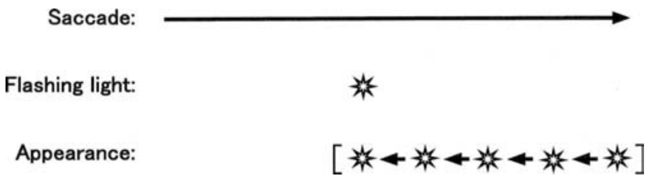


Fig. 3.6 Schematic illustration of a phantom array. When a saccade is executed from the left to the right of a flickering light source, it appears to jump suddenly to the saccade goal and then move back to its original position.

characterized the perceptual pattern as a phantom array, fixed in space, with the sequential order of the flashes creating an impression of apparent motion within the array.

A point light source flickering on and off during a saccade projects a pattern onto the retina, but the perceived pattern differs from the retinal pattern. This difference is thought to reflect the perisaccadic shift of spatiotopic coordinates (local signs) of the retina via a neural signal representing eye position (i.e., EPS).

The damped-EPS model predicts that the flickering stimulus should be perceived moving in accordance with the time course of localization error shown in Fig. 3.1, but Hershberger's experiments indicated that this is not the case. In particular, the finding that the stimulus did not appear to run in the direction of the saccade conflicts with the damped-EPS model. Therefore, Hershberger argued that the findings indicate that EPS does not change gradually but discretely in advance of saccade onsets. According to Hershberger's explanation, the time course of localization error changes gradually is not because EPS is damped but because the time when EPS jumps varies between trials. Consequently, the time course of mean localization error results in a smooth curve. Furthermore, Jordan and Hershberger (1994) pointed out that the discrete jump of the light source in the saccade direction is similar to observations that subjects with extraocular paralysis reported seeing the world jump in the direction of the intended eye movement (Stevens et al. 1976) and the findings that a foveal afterimage appears to displace discretely at the time of saccade execution (Grusser et al. 1987).

3.4.2 Temporal factors

Sogo and Osaka (2001) argued against this idea. If Hershberger's explanation is correct, localization errors of each trial must distribute around zero and the value equal to the saccade amplitude. This is not the case; as shown by many studies, the distribution of localization error is not bimodal before the saccade onset. In addition, Sogo and Osaka found that a phantom array is observed only when a flickering light is presented during saccades, not when presented before saccades. Watanabe et al. (2005) made a similar but more detailed examination. In their experiments, a flickering stimulus of various durations was presented at various points in time around a horizontal saccade. Subjects were required to report the positions of the right and left ends of phantom arrays they perceived. The results were clear: a phantom array is perceived only when the flickering stimulus overlaps in time with saccade execution. Watanabe et al. also stressed that the perceived lengths and positions for presaccadic flickers do not coincide with the prediction from the time course of localization for a perisaccadic single flash, and rejected simple cancellation theory as an explanation of localization of perisaccadic continuous flickers.

3.4.3 Issues surrounding the phantom array

It is still unclear how to explain the discrepancy in perisaccadic localization errors reported for a single flashed target and a successively flashed target. O'Regan (1984) projected seven flashes at consecutive 5-msec intervals at the same retinal positions of the subject's eye during a 12 deg saccade. The subjects perceived a small clump of flashes in which the successive flashes were superimposed to a greater or lesser extent. O'Regan interpreted the results as showing that the EPS grows no more than one-sixth of the total saccade extent. Furthermore, the results suggest the possibility that perisaccadic stimuli are localized at a fixed position in space as long as they are projected successively to the same retinal position, whereas they are seen at various positions when projected to different retinal positions. This hypothesis agrees with the observation that a phantom array is observed only when a flashing light source is presented during a saccade and never before or after a saccade. The latter is consistent with the finding that a continuous visual stimulus is perceived at its veridical position as long as it is extinguished before a saccade onset (Schlag & Schlag-Rey 1995; Cai et al. 1997).

A related finding was that when a subject made a horizontal saccade from the left to the right and a target stimulus moved synchronously with the saccade in the upward direction, the target initially appeared to jump to the right (i.e., in the saccade direction) and then moved toward the actual position of the target after the saccade (Mateeff 1978). Kennard et al. (1971) reported similar observations. In their experiments, the stimulus began to move well before or immediately after the saccade onset. In the former case, the illusory trajectory of the spot was similar to the mislocalization curves reported for a perisaccadically flashed light spot; the spot was seen to move toward the saccade goal before the eye starts to move, and abrupt jumping of targets reported by Mateeff (1978) was not observed. On the other hand, when the stimulus began to move immediately after the saccade onset, it appeared abruptly at the position near the saccade goal and then moved in the direction opposite to the saccade. Thus, abrupt jumping of targets toward a saccade goal seems to occur only when the target begins to move at the same time as the saccade onset.

Strictly speaking, however, things are more complicated. We recently conducted experiments in which ten vertically arranged LEDs were used instead of a light beam from an oscilloscope (Honda 2006). The subjects were required to make 18 deg horizontal saccades from left to right across the array of LEDs. At the time of the saccade, the LEDs flashed one by one for 2 msec from top to bottom. The ISI between each LED flashing was 2 msec, and therefore it took 38 msec from the first LEDs on to the last LEDs off. The distance between the top LED and the bottom LED was 18 deg. The subjects reported their perceptions by drawing what they saw, that is, a trajectory of the successively flashed LEDs. The results were impressive. When the first LED was flashed simultaneously with the saccade onset, the LED appeared at the right of its actual position and then moved diagonally toward the lower left position. The same rightward displacement of the first LED was also observed when it was flashed immediately before or after the saccade onset, as was the case when a single target was flashed perisaccadically in the dark. Therefore, it is evident that EPS is involved in generating the illusory trajectory of a perisaccadically flashed array of LEDs. Our findings suggest that the abrupt jumping of targets toward a saccade goal observed in Hershberger's phantom array experiment is explained by hypothesizing that a flickering stimulus is localized at its veridical position as long as it stimulates the same position on the retina. However, to verify this hypothesis, further research is needed, and it is desirable to explain without contradiction the discrepancy between perisaccadic localization errors reported for a single flashed target and those for a successively flashed target (a phantom array).

Acknowledgments

This work was supported by Grant-in-Aid for Scientific Research (KAKENHI #16530466) from JSPS and a Grant for Promotion of Niigata University Research Projects.

References

Awater, H., & Lappe, M. (2004). Perception of visual space at the time of pro- and anti-saccades. *J Neurophysiol* **91**: 2457–2464.

Bischof, N., & Kramer, E. (1968). Untersuchungen und Uberlegungen zur Richtungswarnehmung bei willkurlichen sakkadischen Augenbewegungen. *Psychologische Forschung* **32**: 185–218.

Bockisch, C. J., & Miller, J. M. (1999). Different motor systems use similar damped extraretinal eye position information. *Vision Res* **39**: 1025–1038.

Boucher, L., Groh, J. M., & Hughes, H. C. (2001). Afferent delays and the mislocalization of perisaccadic stimuli. *Vision Res* **41**: 2631–2644.

Brenner, E., Meijer, W. J., & Cornelissen, F. W. (2005). Judging relative positions across saccades. *Vision Res* **45**: 1587–1602.

Burr, D. C., Morrone, C., & Ross, J. (2001). Separate visual representations for perception and action revealed by saccadic eye movements. *Curr Biol* **11**: 798–802.

Cai, R. H., Pouget, A., Schlag-Rey, M., & Schlag, J. (1997). Perceived geometrical relationships affected by eye-movement signals. *Nature* **386**: 601–604.

Cho, S., & Lee, C. (2003). Expansion of visual space after saccadic eye movements. *J Vis* **3**: 906–918.

Currie, B., McConkie, G., Carlson-Radvansky, L., & Irwin, D. (2000). The role of the saccade target objects in the perception of a visually stable world. *Atten Percep Psychophys* **62**: 673–683.

Dassonville, P., Schlag, J., & Schlag-Rey, M. (1992). Oculomotor localization relies on a damped representation of saccadic eye displacement in human and nonhuman primates. *Vis Neurosci* **9**: 261–269.

Dassonville, P., Schlag, J., & Schlag-Rey, M. (1995). The use of egocentric and exocentric location cues in saccadic programming. *Vision Res* **35**: 2129–2199.

Deubel, H., Bridgeman, B., & Schneider, W. X. (1998). Immediate post-saccadic information mediates space constancy. *Vision Res* **38**: 3147–3159.

Duhamel, J. R., Colby, C. L., & Goldberg, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science* **255**: 90–92.

Grusser, O.-J., Krizic, A., & Weiss, L. (1987). Afterimage movement during saccades in the dark. *Vision Res* **27**: 215–226.

Hallett, P. E. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Res* **18**: 1279–1296.

Hallett, P. E., & Lightstone, A. D. (1976a). Saccadic eye movements towards stimuli triggered by prior saccades. *Vision Res* **16**: 99–106.

Hallett, P. E., & Lightstone, A. D. (1976b). Saccadic eye movements to flashed targets. *Vision Res* **16**: 107–114.

Hansen, R. M., & Skavenski, A. A. (1977). Accuracy of eye position information for motor control. *Vision Res* **17**: 919–926.

Hansen, R. M., & Skavenski, A. A. (1985). Accuracy of spatial localization near the time of saccadic eye movements. *Vision Res* **25**: 1077–1082.

von Helmholtz, H. (1866). *Handbuch der Physiologischen Optik*. Leipzig: Voss.

Hershberger, W. (1987). Saccadic eye movements and the perception of visual direction. *Perception & Psychophysics* **41**: 34.

Honda, H. (1989). Perceptual localization of visual stimuli flashed during saccades. *Perception & Psychophysics* **45**: 162–174.

Honda, H. (1990). Eye movements to a visual stimulus flashed before, during, or after a saccade. In M. Jeannerod (ed.), *Attention and Performance XIII: Motor Representation and Control* (567–582). Hillsdale, NJ: Erlbaum.

Honda, H. (1991). The time courses of visual mislocalization and of extraretinal eye position signals at the time of vertical saccades. *Vision Res* **31**: 1915–1921.

Honda, H. (1993). Saccade-contingent displacement of the apparent position of visual stimuli flashed on a dimly illuminated structured background. *Vision Res* **33**: 709–716.

Honda, H. (1995). Visual mislocalization produced by a rapid image displacement on the retina: Examination by means of dichoptic presentation of a target and its background scene. *Vision Res* **35**: 3021–3028.

Honda, H. (1997). Interaction of extraretinal eye position signals in a double-step saccade task: psychophysical estimation. *Exp Brain Res* **113**: 327–336.

Honda, H. (1999). Modification of saccade-contingent visual mislocalization by the presence of a visual frame of reference. *Vision Res* **39**: 51–57.

Honda, H. (2006). Achievement of transsaccadic visual stability using presaccadic and postsaccadic visual information. *Vision Res* **46**: 3483–3493.

Jordan, J. S., & Hershberger, W. A. (1994). Timing the shift in retinal local signs that accompanies a saccadic eye movement. *Perception & Psychophysics* **55**: 657–666.

Kaiser, M., & Lappe, M. (2004). Perisaccadic mislocalization orthogonal to saccade direction. *Neuron* **41**: 293–300.

Kennard, D. W., Hartmann, R. W., Kraft, D. P., & Glaser, G. H. (1971). Brief conceptual (nonreal) events during eye movements. *Biol Psychiatry* **3**: 205–215.

Krekelberg, B., Kubischik, M., Hoffmann, K.-P., & Bremmer, F. (2003). Neural correlates of visual localization and perisaccadic mislocalization. *Neuron* **37**: 537–545.

Lappe, M., Awater, H., & Krekelberg, B. (2000). Postsaccadic visual references generated presaccadic compression of space. *Nature* **403**: 892–894.

Mateeff, S. (1978). Saccadic eye movements and localization of visual stimuli. *Perception & Psychophysics* **24**: 215–224.

Matin, L., Matin, E., & Pearce, D. G. (1969). Visual perception of direction when voluntary saccades occur: I. Relation of visual direction of a fixation target extinguished before a saccade to a flash presented during the saccade. *Perception & Psychophysics* **5**: 65–80.

Matin, L., Matin, E., & Pola, J. (1970). Visual perception of direction when voluntary saccades occur: II. Relation of visual direction of a fixation target extinguished before a saccade to a subsequent test flash presented before the saccade. *Perception & Psychophysics* **8**: 9–14.

Matin, L., & Pearce, D. G. (1965). Visual perception of direction for stimuli flashed during voluntary saccadic eye movements. *Science* **148**: 1485–1488.

Matsumiya, K., & Uchikawa, K. (2001). Apparent size of an object remains uncompressed during presaccadic compression of visual space. *Vision Res* **41**: 3039–3050.

Miller, J. M. (1996). Egocentric localization of a perisaccadic flash by manual pointing. *Vision Res* **36**: 837–851.

Milner, A. D., & Goodale, M. A. (1995). *The Visual Brain in Action*. Oxford: Oxford University Press.

O'Regan, K. (1984). Retinal versus extraretinal influences in flash localization during saccadic eye movements in the presence of a visible background. *Perception & Psychophysics* **36**: 1–14.

Ross, J., Morrone, M. C., & Burr, D. C. (1997). Compression of visual space before saccades. *Nature* **386**: 598–601.

Schlag, J., & Schlag-Rey, M. (1995). Illusory localization of stimuli flashed in the dark before saccades. *Vision Res* **35**: 2347–2357.

Sogo, H., & Osaka, N. (2001). Perception of relation of stimuli locations successively flashed before saccade. *Vision Res* **41**: 935–942.

Sommer, M. A., & Wurtz, R. H. (2002). A pathway in primate brain for internal monitoring of movements. *Science* **296**: 1480–1482.

Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic responses produced by visual inversion. *J Comp Physiol Psychol* **43**: 482–489.

Stevens, J. K., Emerson, R. C., Gerstein, G. L., Kallos, T., Neufeld, G. R., Nichols, C. W., et al. (1976). Paralysis of the awake human: visual perceptions. *Vision Res* **15**: 93–98.

Tolias, A. S., Moore, T., Smirnakis, S. M., Tehovnik, E. J., Siapas, A. G., & Schiller, P. H. (2001). Eye movements modulate visual receptive fields of V4 neurons. *Neuron* **29**: 757–767.

Umeno, M. M., & Goldberg, M. E. (1997). Spatial processing in the monkey frontal eye field. I. Predictive visual responses. *J Neurophysiol* **78**: 1373–1383.

Von Holst, E., & Mittelstaedt, H. (1954). Das Reafferenzprinzip. *Naturewissenschafften* **37**: 464–476.

Walker, M. F., Fitzgibbon, E. J., & Goldberg, M. E. (1995). Neurons in the monkey superior colliculus predict the visual result of impending saccadic eye movements. *J Neurophysiol* **73**: 1988–2003.

Watanabe, J., Noritake, A., Maeda, T., Tachi, A., & Nishida, S. (2005). Perisaccadic perception of continuous flickers. *Vision Res* **45**: 413–430.

4

Visual and nonvisual factors in perisaccadic compression of space

MARKUS LAPPE, LARS MICHELS, AND HOLGER AWATER

Summary

The perceptual stability of visual space becomes fragile in the wake of a saccadic eye movement. Objects flashed shortly before a saccade are mislocalized toward the saccade target. Traditional accounts for this effect have associated the mislocalizations with sluggishness of the efference copy signal, which is important in space perception across eye movements. Recent theories of space perception, however, have emphasized a role for visual memory in the generation of transsaccadic spatial stability. We have investigated the role of visual processes and their interactions with efference copy signals in the perisaccadic compression of space. In our experiments, subjects performed saccades in front of a computer display while visual stimuli were briefly flashed on the screen just before or during the saccade. Subjects had to report the perceived location of the flash. When the saccade target's position was visibly available after the saccade, the perceived location of the flash was compressed toward the target's position. This compression occurred not only along the axis of the saccade but also for parts of visual space along a direction orthogonal to the saccade. When the saccade target was not visibly available after the saccade, the perceived location of the flash showed only a slight shift in saccade direction. In this condition, however, the perceived location of the saccade target was drawn toward the position of the flash. We propose a framework that consists of pre- and postsaccadic processes to explain these findings.

4.1 Introduction

The overtly observable aspects of a saccade are small. Saccades involve only a tiny movement of the eyeball. Yet, this tiny motion is orchestrated by coordinated activity in a large network of interconnected brain areas. It spans from the brain stem to midbrain and cerebellar areas and several parts of the cortex. Moreover, the tiny eye movement of a saccade exerts massive influences on visual perception. On the one hand, it provides necessary updates for vision to scan a scene, while on the other hand it causes dramatic changes to the visual input, giving rise to problems of visual stability. In this article, we will discuss effects that saccades induce in the localization of visual objects around the time the saccade is initiated.

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

It has long been known that visual localization errors occur when a stimulus is briefly presented in the wake of a saccade (Matin & Pearce 1965; Bischof & Kramer 1968; Honda 1989; Dassonville et al. 1995). In these experiments visual probe stimuli, usually small spots of light or luminous bars oriented orthogonal to saccade direction, are presented for a few milliseconds as an observer prepares and executes a saccade. The observer is instructed to report the apparent position of the probe either by using a pointing device that can be adjusted after the saccade or by relating the probe to a reference stimulus such as a ruler or a previous or subsequent test stimulus. Localization is found to be close to veridical for probe stimuli presented more than 100 msec before and after a saccade. Stimuli presented from about 100 msec before to sometime during the saccade are mislocalized. The spatial and temporal pattern of localization errors obtained in such experiments informs discussions about the origin of the mislocalization and the brain processes that occur during the preparation and execution of saccades.

A number of studies have reported mislocalizations that consist of a spatially uniform shift of all apparent positions. This shift is in the direction of the saccade from up to 100 msec before saccade onset until about saccade onset (Matin et al. 1970; Honda 1991; Dassonville et al. 1995). During the saccade, the shift is against saccade direction but still spatially uniform (Matin et al. 1969; Honda 1991; Dassonville et al.). This pattern of errors may be explained if the presumed reference signal of eye position (von Helmholtz 1896; von Holst & Mittelstaedt 1950; Sperry 1950) does not correctly match the time course of the true eye position (Matin et al. 1969; Honda 1991; Schlag & Schlag-Rey 2002) or if the latency and persistence of the visual response is taken into account in the combination of visual and eye position signals (Matin et al. 1969; Pola 2004). Other studies have reported errors that are spatially nonuniform and consist of an apparent compression of spatial positions so that they cluster around the saccade target (Bischof & Kramer 1968; Morrone et al. 1997; Ross et al. 1997). These mislocalizations begin about 50 msec before saccade onset and peak close to saccade onset.

In the following sections we discuss several experiments that investigated the role of visual and nonvisual factors in perisaccadic compression, present new data from an experiment investigating potential motor contributions from subsequent saccades, and propose a conceptual framework to explain the various findings. For a detailed discussion of possible mechanisms of the uniform perisaccadic shift, see the chapters by Matin and Honda in this volume.

4.2 Visual factors

Visual factors play an important role in distinguishing shift and compression. Studies that reported spatially uniform shifts were typically conducted in darkened labs (Honda 1989; Dassonville et al. 1995; Cai et al. 1997). At the onset of a saccade, only the probe stimulus was visible with no other visual object that could serve as a positional reference. On the other hand, studies that found compression involved stimulus presentation on a screen that contained, next to the brief probe flash, reference objects such as a visual background or a

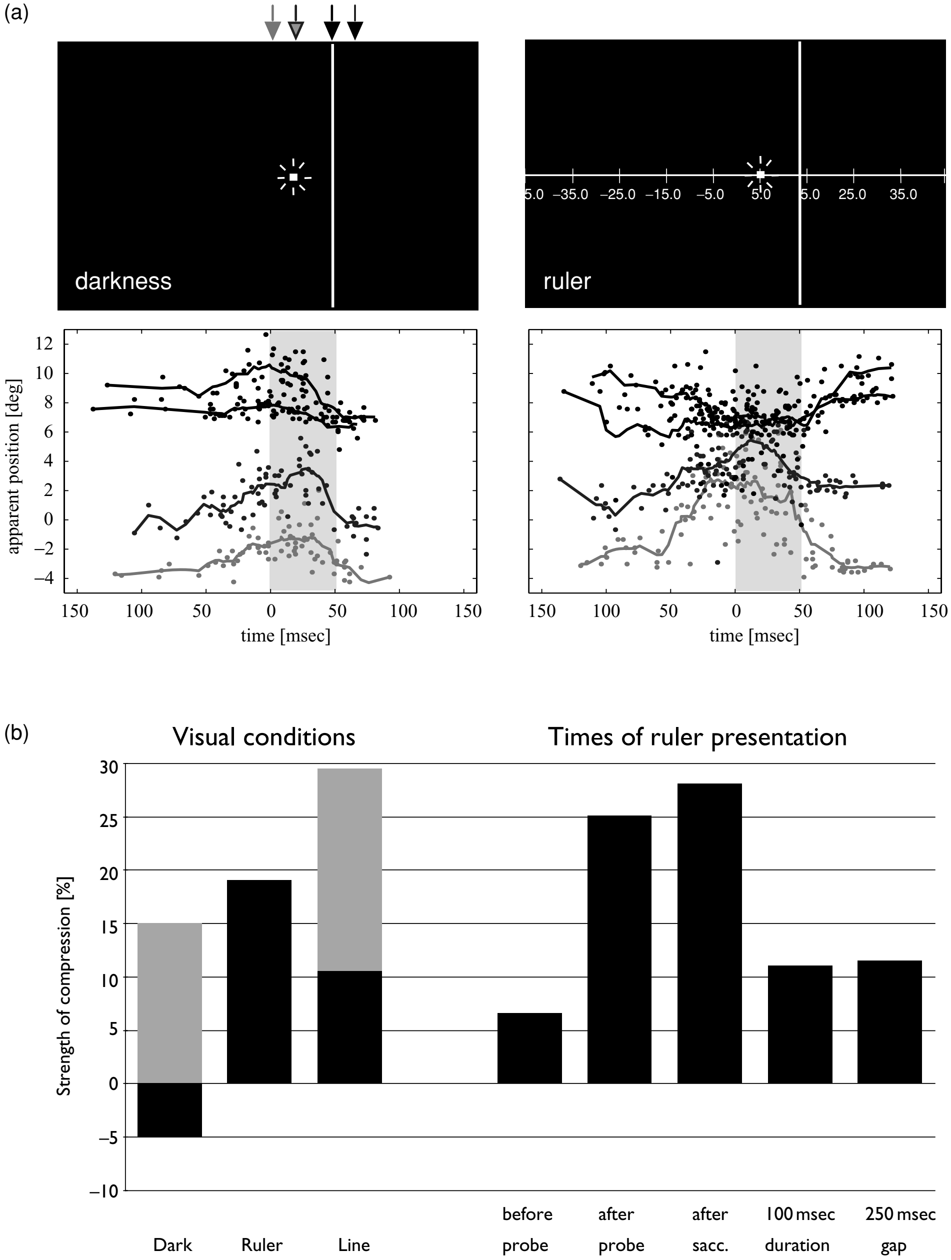


Fig. 4.1 The dependence of perisaccadic compression on visual references. (a) Different patterns of perisaccadic mislocalization observed in the same subject in darkness (left) and with a ruler identifying the target position after the saccade (right). Probes (flashed bars) were presented randomly at one of four locations (at −2.6, 2.6, 10, and 13 deg). Apparent positions at various times relative to saccade onset are shown by individual dots. Lines give running averages through the data. Around saccade

ruler to which the observer had to relate the probe position (Bischof & Kramer 1968; Ross et al. 1997; Morrone et al. 1997; see also Honda 1993). Testing the same observers in both conditions we found that the presence or absence of a ruler on the stimulus screen strongly influenced the observed pattern of mislocalization (Lappe et al. 2000). In one experiment, observers performed a visually guided saccade in front of a large dark projection screen in a dark experiment room. They initially fixated a bright fixation point 6.4 deg left of straight ahead that jumped to a position 6.4 deg to the right of straight ahead (12.8-deg saccade amplitude) and was extinguished 50 msec later. Because the average latency of the saccade that followed the jump was around 200 msec, no visual reference stimuli were available at the time of saccade onset or afterward. The probe was a vertical bar flashed at one of four locations around the saccade target (at –2.6, 2.6, 10, and 13 deg relative to straight ahead). The observer reported the perceived probe position by adjusting a mouse pointer appearing 500 msec after the saccade. The apparent position of those probes was analyzed in relation to their presentation time relative to saccade onset. We found that the four probe positions were uniformly shifted first in and then against the direction of the saccade similar to the shift reported in earlier studies (Fig. 4.1(a), left). However, when we introduced, in a second condition, a continuously visible ruler on the screen, which could serve as a visual reference for probe and target position, the spatial pattern of mislocalization changed and a strong compression of the apparent probe position toward the saccade target was observed (Fig. 4.1(a), right). To assess the strength of compression quantitatively, we calculated the standard deviation between the four apparent probe positions at each point in time for a measure of the average apparent separation of the probe positions. We then calculated the mean apparent separation in the 50 msec before saccade onset, that is, when the eye was still fixating, and compared it to the mean separation 100 msec before and after the saccade. This resulted in a percent measure of the compression immediately before the saccade. Based on this measure there was no compression in the darkness condition and strong compression in the ruler condition (Fig. 4.1(b), dark bars). We concluded that the strength of compression was dependent on the presence of the ruler.

The ruler not only provided spatial references about probe and target positions but also retinal illumination that might have modified the contrast of the probe. In further experiments, we found that both factors influence the strength of compression. The influence

Fig. 4.1 (*cont.*)
onset, apparent positions are shifted in saccade direction in the dark condition and compressed onto the saccade target location (at 6.4 deg) in the ruler condition. (b) Strength of perisaccadic compression (dark bars) measured as the ratio of the mean apparent separation of bars within 50 msec before saccade onset and the mean apparent separation 100 msec before and after the saccade for various visual conditions and presentation times of the ruler. The light bars in the dark and line conditions show the strength of compression when instead of the apparent probe positions the apparent distance between the probe and the saccade target was calculated. These values differ because in the dark and line conditions the saccade target is mislocalized toward the location of the probe (Awater & Lappe 2006). The combined effect of both mislocalizations leads to increased compression.

of contrast was measured in an experiment in which the contrast of the probe to the background was varied (Michels & Lappe 2004). In this experiment, the screen background was dark gray (luminance 13.2 cd/m^2), whereas the probes had luminance between 14.3 cd/m^2 and 61.3 cd/m^2. The strength of apparent compression varied with stimulus contrast such that the strongest compression was observed at the lowest contrast.

We measured the influence of visual references by varying when the ruler was present on the screen (Lappe et al. 2000). We found that compression was weak when the ruler was present before, but not after, the flash of the probe (Fig. 4.1(b), before and after probe conditions). Compression was strong when the ruler was present after the saccade, independent of whether it was visible at the time the probe was flashed (Fig. 4.1(b), ruler and after-saccade conditions). This suggests that the spatial reference information provided by the ruler is evaluated mainly after the saccade and that the apparent compression of stimuli presented before the saccade is at least partly dependent on signal processing that occurs after the saccade. To further determine the time course over which the ruler influences localization errors, we varied the onset time and duration of the ruler presentation after the saccade. When the ruler was turned off when the probe flashed and switched on only 250 msec later, the apparent compression was diminished (Fig. 4.1(b), 250 msec-gap condition). When the ruler was present only from saccade offset up to 100 msec later, compression strength was also diminished (Fig. 4.1(b), 100 msec-duration condition). These results suggest that there is a temporal window immediately after the saccade when the presence of visual references influences strength of compression.

To compare the contribution of retinal stimulation and contrast reduction provided by the ruler with that of visual references provided by the ruler we introduced a further condition. In this condition, the ruler was replaced by a horizontal white line, visually identical to the ruler but with no tick marks or numbers present (Awater & Lappe 2006). This line provides retinal stimulation and contrast reduction but no visual references along the saccade direction. Consistent with the prediction that both factors contribute to compression, the strength of compression was about half of that observed with the tick marks (Fig. 4.1(b), line condition compared to ruler condition). But a second perceptual effect became apparent in this experiment: although the saccade target was always presented at the same spatial location, it appeared to be at different locations for the different probe positions. Asking observers to report the apparent location of the saccade target rather than of the probe quantitatively confirmed this observation (Awater & Lappe 2006). Depending on the probe position the saccade target could be mislocalized by several degrees of visual angle. This mislocalization was such that the distance between the apparent target position and the apparent probe position was compressed. When we calculated a compression index for the apparent probe–target distance, a sizable compression became apparent in the darkness condition and the compression in the line condition became equal to the compression in the ruler condition (Fig. 4.1(b), gray bars). Although the flash of the probe influenced perceived target locations, saccadic amplitudes remained veridical when the probe was flashed within 100 msec before saccade onset. Thus, the saccade target's apparent position, which was visually indicated only for 50 msec after fixation point onset, could be decoupled from the motor command that drives the saccade. Moreover, the mislocalization of the saccade target occurred also without a saccade (Awater & Lappe 2006). When the sequence of fixation point offset, target on- and offset, and probe presentation was run with the explicit instruction to keep fixation throughout the trial (i.e., not make a saccade to the target), the perceived location of the target was shifted toward the probe position much like in the saccade condition. Compression of the probe positions did not occur in this situation.

The apparent position shift of the target in saccade and fixation conditions suggests that the target shift results from visual factors unrelated to the saccade, or possibly from covert saccade planning that might have occurred in the fixation condition. We will now turn to the role of motor parameters in perisaccadic compression and return to the target mislocalization in the presentation of the proposed framework for perisaccadic mislocalization.

4.3 Motor factors

The necessity of saccade execution for the compression effect has been investigated by comparing perception during executed saccades with perception during "simulated" saccades, in which the visual stimulation during the saccade is mimicked as closely as possible by moving the visual image at saccadic speeds while the observer fixates. Apparent position shifts (Morrone et al. 1997) and compression (Ostendorf et al. 2006) of probe stimuli also occur in such experiments, but the reported compression is not as strong as in experiments with real saccades. Thus, the execution of the saccade likely contributes to the compression. Because perisaccadic mislocalization during real saccades begins already for stimuli presented 50–100 msec before the saccade, that is, when the eye is still directed toward the fixation point, the role of the actual movement of the eye is not clear. Motor signals that drive the eye movement – or corollaries of these signals – are likely to be involved. Two experimental paradigms, the antisaccade task and saccadic adaptation, have been used to differentiate between motor execution (or planning) of the saccade and visual stimulation factors such as the saccade target.

In the antisaccade task, when the fixation point jumps in a certain direction, observers are instructed to withhold the saccade in the direction of the jump and to perform a saccade in the opposite direction of the jump. In this case, the peripheral visual stimulus, that is, the saccade target in regular or prosaccade trials, is decoupled from the motor planning and execution of the saccade. Thus, one may ask whether perisaccadic compression is directed toward the visual image of the saccade target or toward the motor signal of the saccade. In these experiments, perisaccadic localization errors were directed toward the actual saccade, not toward the visual target stimulus (Awater & Lappe 2004). Strength of compression was identical in anti- and prosaccade trials. This suggests that the actual motor plan underlies the compression.

In the saccade adaptation paradigm, the visual target location and the motor execution of the saccade become decoupled through an adaptation process. As the observer executes a saccade induced by a target jump of a defined size the target is slightly displaced during the saccade. Therefore, at the end of the saccade there is an error between the landing position of the eye and the postsaccadic target position. When this procedure is repeated for a number of trials the amplitude of the induced saccade gradually adapts such that the saccade will end closer to the postsaccadic target position. After the adaptation procedure, a target jump of the defined size induces a saccade of the adapted amplitude. Thus, target location and motor execution are decoupled. Awater et al. (2004) measured perisaccadic compression after saccadic adaptation. They presented probe flashes around the occurrence of adapted saccades. Adaptation changed the pattern of compression such that probe positions at saccade onset were seen closer to the landing point of the eye rather than the initial target location. Like the results of the antisaccade study, the results of the adaptation experiment suggest that the motor plan of the actually executed saccade is the driving force behind the compression. There is, however, a caveat to this conclusion because the experiments also showed that probes outside the temporal range of perisaccadic compression, that is, probes presented more than 100 msec before saccade onset, were mislocalized. This early presaccadic mislocalization shifted apparent probe positions in the direction of adaptation. For instance, when back-stepping the target during the saccade reduced the saccadic amplitude, perceived presaccadic probe positions similarly shifted backward. Thus, the clustering of apparent probe positions near the actual saccade landing position at saccade onset may result partly from a combination of adaptation-induced position shifts and perisaccadic compression toward the initial target position.

The previous studies show that the saccadic motor plan, or a corollary of it, is involved in generating the compression. An oculomotor feedback signal of the saccadic motor plan may originate from a number of brain structures involved in saccade planning or execution. Depending on where the signal originates in the brain it might be encoded as direction and amplitude or as a two-dimensional map of saccade goal position. The second, but not the first, case predicts that mislocalizations should occur in a two-dimensional manner around the saccade goal position. We tested this in an experiment with probes of small light dots arranged in a grid around the saccade target position. The perisaccadic mislocalization of these probe positions showed a clear two-dimensional pattern such that positions at high eccentricities were mislocalized in oblique directions toward the saccade goal (Kaiser & Lappe 2004). Thus, the saccade-related signal that drives the mislocalization is a signal of goal position rather than saccade amplitude.

If compression is induced by an oculomotor feedback signal, then one may also ask how compression is related to the planning and execution of subsequent saccades. It is known that perceptual and motor processes in single tasks can be decoupled. Specifically for saccadic compression, it has been reported that pointing movements toward the apparent position of the perisaccadic probe do not exhibit compression, although perceptual judgments of probe position do (Burr et al. 2001). We were therefore interested in the question of whether targeting saccades to the apparent probe position, conducted after the primary saccade to the target, shows evidence of compression. In these experiments, subjects were seated in front of a monitor that displayed an initial fixation point 10 deg left of center against a gray visual background containing a black ruler. When the fixation point jumped 20 deg to the right (toward position 10 deg) subjects performed a saccade toward that position.

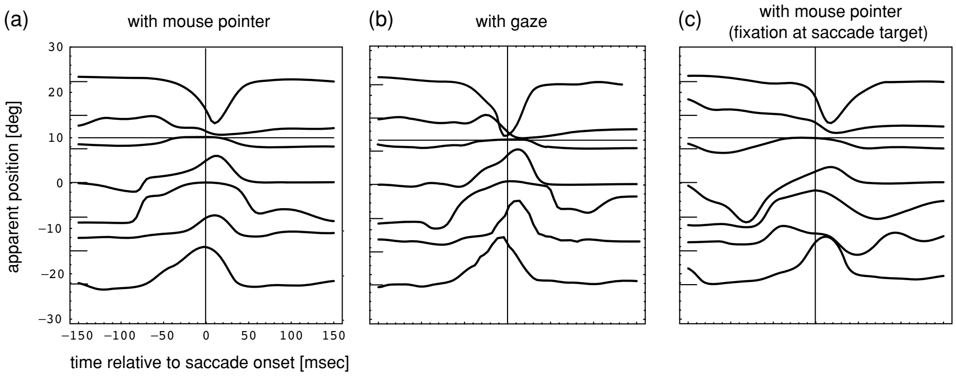


Fig. 4.2 Perisaccadic localization quantified by different reporting procedures in a single subject. The saccade was executed from position −10 deg to position 10 deg with respect to straight ahead. Probes were flashed at seven positions (−22.4 deg, −14.9 deg, −7.6 deg, 0 deg, 7.6 deg, 14.9 deg, 22.4 deg) in randomized order. (a) Apparent positions (running averages, individual data points not shown) indicated by a mouse pointer that appeared 500 msec after the saccade and the subject moved to the perceived probe location. The subject was free to move his eyes at this time. (b) The subject was instructed to first make a saccade to the target followed by a saccade to the probe that was flashed around the occurrence of the first saccade. Apparent probe positions were computed from the landing position of the second saccade. (c) Apparent probe positions indicated by mouse pointing as in (a). However, in this condition the subject was required to maintain fixation on the first saccade target while adjusting the mouse pointer using peripheral vision. The compression was similar in all three conditions and for two additional subjects.

A vertical bar was flashed as the probe for 8 msec at a random time around saccade onset at one of seven possible locations (−22.4 deg, −14.9 deg, −7.6 deg, 0 deg, 7.6 deg, 14.9 deg, 22.4 deg). Subjects were instructed to direct their gaze after the primary saccade to the apparent position of the probe stimulus and to keep fixating that position. Gaze direction was measured with an SMI EyeLink video-based eye tracker at a temporal resolution of 250 Hz. Final gaze positions after the secondary saccade were taken as measurements of the targeting saccade to the apparent probe positions. These measurements were compared to perceptual ratings taken in separate trials in which subjects had to indicate the apparent probe position with a mouse pointer. Figure 4.2(a) and (b) shows results of this experiment for one subject. The lines give running averages of apparent probe positions for the seven true probe positions in the 300 msec around the saccade. Saccade onset is at 0 msec. A comparison of the plots for mouse pointing (Fig. 4.2(a)) and gaze direction reveals that compression is similar in the two conditions. The percentage of compression in the 50 msec before the saccade was the same in both conditions. This shows that perceptual and saccadic mislocalizations are identical. Thus, the compression observed in the perceptual judgments is reflected also in the positions of targeting saccades. Similar results have been reported in investigations of the perisaccadic shift of apparent positions in the dark using the double-step saccade paradigm (Dassonville et al. 1995; Honda 1997).

Because mislocalizations in perceptual judgments and targeting saccades are so similar, one must ask whether perceptual judgments rely on the execution of subsequent saccades. This would seem possible if subjects, after the primary saccade, conduct first a saccade to the apparent probe position before they report the probe position verbally or with a pointer. We therefore tested a further condition in which subjects had to give perceptual judgments with a mouse pointer but were required to keep fixation at the target position after the primary saccade (Fig. 4.2(c)). Compression in this condition was of equal strength to the compression observed under free eye movements after the primary saccade and under conditions of targeting saccades after the primary saccade. Thus, we conclude that secondary targeting saccades are not necessary for compression to occur and that the compression observed in the targeting saccades likely results from the use of the compressed perceptual signal for planning these eye movements.

4.4 A conceptual framework for perisaccadic mislocalization

The studies described previously revealed different types of perceptual mislocalization at the time of saccadic eye movements. In some conditions, a uniform mislocalization in the direction of the saccade occurred. In other conditions, mislocalizations resembled a compression of visual space around the saccade target. The uniform shift occurred when no visual references were available immediately after the saccade and when probe stimuli were of high contrast. Compression occurred when probe stimuli were of low contrast and visual references were present immediately after the saccade. Thus, the visual parameters of the stimuli as well as the presence or absence of visual references immediately after the saccade are important for the compression of the apparent position of perisaccadically flashed objects.

An explanation for the described properties of the compression requires the combination of pre- and postsaccadic processes, both involving visual and nonvisual factors (Fig. 4.3). Visual information about the target position and the probe position is registered before the saccade and encoded in memory. After the saccade, the memory representation has to be accessed and combined with new visual or nonvisual information about the current eye position. Because neither the target nor the probe is visible after the saccade, both positions have to be reconstructed from memory signals and the available postsaccadic information. Our proposal rests on two assumptions. The first assumption is that the apparent distance between the saccade target and the probe is compressed in the encoding of the presaccadic scene. The second assumption is that the reconstruction of the scene after the saccade prefers postsaccadic visual reference information to nonvisual eye position signals. In the following we will discuss the support for these assumptions provided by the studies presented previously.

The role of postsaccadic visual references can be understood by looking at studies of transsaccadic visual stability. Current ideas on visual stability across saccades hold that most of the visual information before the saccade is discarded, and a new representation of the visual world is generated from new incoming information after the saccade (Bridgeman

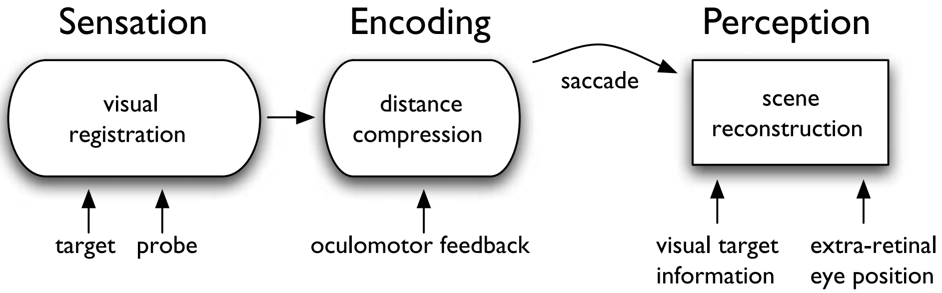


Fig. 4.3 Schematic illustration of the steps involved in perisaccadic compression. First, the visual signals of the target and the probe have to be registered. The location of the target and probe along with the distance between the two are encoded for transsaccadic memory. Either the registration or the encoding in memory is subject to a compression of the distance between the target and the probe. Presumably this compression reflects the action of a nonvisual modulatory oculomotor feedback signal. This is assumed to occur before the saccade. After the saccade, the layout of the presaccadic scene has to be reconstructed from the presaccadic encoding and the available postsacccadic information about target position (visual) and eye position (nonvisual). Because the system prefers to use visual reference information, the saccade target is localized in the visually indicated position, and the probe appears compressed according to the compressed presaccadic memory representation. In the absence of visual target information, nonvisual information about eye position is used for the reconstruction of the presaccadic probe locations. In this case, the apparent probe location is shifted along saccade direction because of errors in the eye position signal.

et al. 1994; Hamker 2005a). Visual stability is realized by a store-and-compare mechanism that uses only limited information from before the saccade (Deubel et al. 1996, 1998; McConkie & Currie 1996; Currie & McConkie 2000). The basic assumption in this theory is that the perceptual world remains stable if nothing changes during the saccade. Thus, the store-and-compare mechanism uses presaccadic information from only a small number of objects, most importantly the saccade goal itself. This information is stored in a non-retinotopic memory across the saccade. Immediately after the saccade the visual system searches the saccade target in a restricted spatiotemporal window and compares the stored representation with the new visual information after the saccade. If the "new" saccade target is identified as the "old" saccade target it will be used for the recalibration of the visual scene. In our experiments, the saccade target was never present after the saccade, but visual reference information from the ruler was available to indicate the position of the saccade target. This reference information might be used instead of the image of the saccade target to localize the saccade's goal position, allowing the reconstruction of visual space from presaccadic memory with respect to the saccade target's ruler location.

The store-and-compare mechanism, thus, could be involved in generating perisaccadic compression. Postsaccadic visual space is constructed from visual information in the post-saccadic scene with limited memory information from the presaccadic scene. This post-saccadic reconstruction is centered on the postsaccadic position of the saccade target. The probes in our experiment are only presented in the presaccadic scene. Therefore, their position has to be reconstructed from visual memory after the saccade. If presaccadic memory mainly preserves information from the saccade target, then it is conceivable that information about other objects is distorted toward the saccade target position. Therefore, we speculate that the observed compression is due either to memory distortions induced by the presaccadic concentration of resources on the saccade target or to distortions within the presaccadic scene.

The observed mislocalization of the saccade target in the white line and darkness conditions supports this view. An apparent shift of the target position toward the probe position occurred both in the presence and in the absence of a saccade. This suggests that the registered and encoded distance between target and probe is compressed independently of the occurrence of a saccade. However, the position of the probe was mislocalized only when the subject performed a saccade, not when the subject maintained fixation. Within the store-and-compare mechanism the difference between these conditions can be explained because the postsaccadic reconstruction of the scene obviously only becomes necessary when a saccade occurs. In this case, the probe and target positions are retrieved from the encoded compressed distance between target and probe position. If the target position is indicated by visual references, then it can be localized veridically, but the retrieved probe position will appear compressed. If, on the other hand, the target position is not indicated by visual references and the probe location is perceived correctly, then the perceived target position will appear compressed toward the probe.

Why is the target position not derived correctly from extraretinal eye position signals after the saccade made in darkness and in the white line condition? If the saccade target's position is not visually indicated after the saccade, then both the target and the probe positions must ultimately be determined from a combination of visual memory with extraretinal eye position information (Matin et al. 1969; Matin 1972). If the task is to indicate the probe, then the probe's position should be determined from its presaccadic retinal location (or memory thereof) and the eye-position signal as proposed earlier (Matin et al. 1969; Honda 1991; Schlag & Schlag-Rey 2002; Pola 2004). This is consistent with our results in the darkness condition. The same procedure may, in principle, be used to locate the target position as well. However, our results on the apparent target position show that this is not the case. Because the target is mislocalized toward the probe, its position must be derived from the probe's position (including extraretinal information) and the memory of the compressed distance between the probe and the target in the presaccadic scene.

The previous considerations may explain why, if we assume a compression of distance between target and probe in the presaccadic encoding, probe positions in the postsaccadic judgment appear compressed toward the target in the presence of visual references but not without visual references. They also explain why the target may appear shifted toward the probe without visual references both after saccades and during fixation. However, they do not explain why the distance between the probe and the target is compressed in the first place. The observations described previously point to a presaccadic origin of this compression. The antisaccade and the saccade adaptation results show that a motor command signal must be involved that may be dissociated from the visual target position. Thus, an oculomotor feedback signal is most likely responsible for the presaccadic induction of compression.

This signal would originate from motor command structures and be fed back to visual areas. Such signals have been described (Sommer & Wurtz 2002; Moore & Armstrong 2003) and linked to the boosting of the sensitivity of visual neurons surrounding the target position to enhance spatial processing in that area (Hamker 2005b). Such a modulatory influence on neural activities may distort the representation of stimulus locations in the map surrounding the saccade target, leading to changes in position signals (Hamker et al. 2008). Suppose that the stimulus excites a population of neurons with partially overlapping receptive fields in a retinotopic cortical map. For some of these neurons, the stimulus will fall onto the receptive field center and the neuron will respond strongly; for other neurons, the stimulus will fall toward the edge of the receptive field and the neuron will respond poorly. The activity in the map will show a peak at the position where the receptive field center falls on the stimulus position. Now assume further that the feedback signal from the eye movement command increases the gain of the neurons with receptive fields close to the saccade target. Those neurons that contain the stimulus in their receptive field *and* that are close to the saccade target will respond stronger than those that contain the stimulus in their receptive field but are further away from the target. Importantly, those neurons close to the saccade target that contain the stimulus at the edge of their receptive field will become more responsive to the stimulus. Thus, in the map, the hill of activity will become asymmetrical and the peak, or at least the center of mass of the activity, will be shifted toward the saccade target. If the perceived position of the stimulus is derived from the distribution of the population activity in such a cortical map (certainly a sensible assumption), then the perceived position must be distorted toward the target. Quantitative analysis of this model shows that it is consistent with the two-dimensional mislocalizations (Kaiser & Lappe 2004) and their similarity to cortical magnification factors (VanRullen 2004; Hamker et al. 2008). It is also consistent with the finding that low-contrast stimuli lead to stronger compression (Michels & Lappe 2004) because the gain modulation is most effective for visual stimuli that yield nonsaturated responses such as stimuli falling on the edge of the receptive field.

4.5 Conclusion

The framework we have proposed has a presacccadic and a postsaccadic component (Fig. 4.3). Before the saccade, the visual signals of the target and the probe are registered and encoded in memory. Either the registration or the encoding in memory is subject to a compression of the distance between the target and the probe, presumably by the action of a modulatory oculomotor feedback signal. Thus, presaccadic processes induce compression. After the saccade, the spatial locations of the presaccadic objects have to be reconstructed from the presaccadic memory in combination with visual reference information about target position and nonvisual information about eye position. The system prefers to use visual reference information when available, locating the saccade target in the visually indicated position and mislocalizing the probe according to the compressed presaccadic memory representation. If visual reference information about the target position is not available after the saccade, the system must use nonvisual information about eye position for the reconstruction of the presaccadic locations. In this case, the apparent probe location is uniformly shifted due to errors in the eye-position signal, but residues of compression in the presaccadic encoding can be seen in the mislocalization of the target position.

Acknowledgments

M.L. is supported by the German Science Foundation DFG LA-952/2 and LA-952/3, the German Federal Ministry of Education and Research Project "Visuo-spatial cognition," and the EC projects Drivsco and Eyeshots.

References

Awater, H., Burr, D., Lappe, M., Morrone, M. C., & Goldberg, M. E. (2004). The effect of saccadic adaptation on the localization of visual targets. *J Neurophysiol* **93**: 3605–3614.

Awater, H., & Lappe, M. (2004). Perception of visual space at the time of pro- and anti-saccades. *J Neurophysiol* **91**: 2457–2464.

Awater, H., & Lappe, M. (2006). Mislocalization of perceived saccade target position induced by perisaccadic visual stimulation. *J Neurosci* **26**: 12–20.

Bischof, N., & Kramer, E. (1968). Untersuchungen und Überlegungen zur Richtungswahrnehmung bei willkürlichen sakkadischen Augenbewegungen. *Psychologische Forschung* **32**: 185–218.

Bridgeman, B., van der Heijden, A. H. C., & Velichkovsky, B. M. (1994). A theory of visual stability across saccadic eye movements. *Behav Brain Sci* **17**: 247–292.

Burr, D. C., Morrone, M. C., & Ross, J. (2001). Separate visual representations for perception and action revealed by saccadic eye movements. *Curr Biol* **11**: 798–802.

Cai, R. H., Pouget, A., Schlag-Rey, M., & Schlag, J. (1997). Perceived geometrical relationships affected by eye-movement signals. *Nature* **386**: 601–604.

Currie, C. B., & McConkie, G. W. (2000). The role of the saccade target object in the perception of a visually stable world. *Perc Psychophys* **62**: 673–683.

Dassonville, P., Schlag, J., & Schlag-Rey, M. (1995). The use of egocentric and exocentric location cues in saccadic programming. *Vision Res* **35**: 2191–2199.

Deubel, H., Bridgeman, B., & Schneider, W. X. (1998). Immediate post-saccadic information mediates space constancy. *Vision Res* **38**: 3147–3159.

Deubel, H., Schneider, W. X., & Bridgeman, B. (1996). Postsaccadic target blanking prevents saccadic suppression of image displacement. *Vision Res* **36**: 985–996.

Hamker, F. H. (2005a). A computational model of visual stability and change detection during eye movements in real world scenes. *Vis Cogn* **12**: 1161–1176.

Hamker, F. H. (2005b). The reentry hypothesis: the putative interaction of the frontal eye field, ventrolateral prefrontal cortex, and areas V4, IT for attention and eye movement. *Cerebral Cortex* **15**: 431–447.

Hamker, F. H., Zirnsak, M., Calow, D., & Lappe, M. (2008). The peri-saccadic perception of objects and space. *PLoS Comp Biol* **4**(2): e31, 1–15.

Honda, H. (1989). Perceptual localization of visual stimuli flashed during saccades. *Percep Psychophys* **45**: 162–174.

Honda, H. (1991). The time course of visual mislocalization and of extraretinal eye position signals at the time of vertical saccades. *Vision Res* **31**: 1915–1921.

Honda, H. (1993). Saccade-contingent displacement of apparent position of visual stimuli flashed on a dimly illuminated structure background. *Vision Res* **33**: 709–716.

Honda, H. (1997). Interaction of extraretinal eye position signals in a double-step saccade task: psychophysical estimation. *Exp Brain Res* **113**: 327–336.

Kaiser, M., & Lappe, M. (2004). Perisaccadic mislocalization orthogonal to saccade direction. *Neuron* **41**: 293–300.

Lappe, M., Awater, H., & Krekelberg, B. (2000). Postsaccadic visual references generate presaccadic compression of space. *Nature* **403**: 892–895.

Matin, L. (1972). Eye movements and perceived visual direction. In D. Jameson & L. Hurvich (eds.), *Handbook of Sensory Physiology (vol. VII/4) Visual Psychophysics* (331–380). Berlin: Springer-Verlag.

Matin, L., & Pearce, D. G. (1965). Visual perception of direction for stimuli during voluntary saccadic eye movements. *Science* **148**: 1485–1488.

Matin, L., Pearce, E., & Pearce, D. G. (1969). Visual perception of direction when voluntary saccades occur: I. Relation of visual directions of a fixation target extinguished before a saccade to a subsequent test flash presented during the saccade. *Percep Psychophys* **5**: 65–80.

Matin, L., Pearce, E., & Pola, J. (1970). Visual perception of direction when voluntary saccades occur: II. Relation of visual direction of a fixation target extinguished before saccade to a subsequent test flash presented before the saccade. *Percep Psychophys* **8**: 9–14.

McConkie, G. W., & Currie, C. B. (1996). Visual stability across saccades while viewing complex pictures. *J Exp Psychol: Hum Percept Perform* **22**: 563–581.

Michels, L., & Lappe, M. (2004). Dependence of saccadic compression and suppression on contrast. *Vision Res* **44**: 2327–2336.

Moore, T., & Armstrong, K. M. (2003). Selective gating of visual signals by microstimulation of frontal cortex. *Nature* **42**: 370–373.

Morrone, M. C., Ross, J., & Burr, D. C. (1997). Apparent position of visual targets during real and simulated saccadic eye movements. *J Neurosci* **17**: 7941–7953.

Ostendorf, F., Fischer, C., Gaymard, B., & Ploner, C. J. (2006). Perisaccadic mislocalization without saccadic eye movements. *Neurosci* **137**(3): 737–745.

Pola, J. (2004). Models of the mechanism underlying perceived location of a perisaccadic flash. *Vision Res* **44**(24): 2799–2813.

Ross, J., Morrone, M. C., & Burr, D. C. (1997). Compression of visual space before saccades. *Nature* **386**: 598–601.

Schlag, J., & Schlag-Rey, M. (2002). Through the eye, slowly: delays and localization errors in the visual system. *Nature Rev Neurosci* **3**: 191–200.

Sommer, M. A., & Wurtz, R. H. (2002). A pathway in primate brain for internal monitoring of movements. *Science* **296**: 1480–1482.

Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *J Comp Physiol Psychol* **43**: 482–489.

VanRullen, R. (2004). A simple translation in cortical log-coordinates may account for the pattern of saccadic localization errors. *Biol Cybern* **91**: 131–137.

von Helmholtz, H. (1896). *Handbuch der Physiologischen Optik*. Hamburg: Leopold Voss.

von Holst, E., & Mittelstaedt, H. (1950). Das Reafferenzprinzip (Wechselwirkung zwischen Zentralnervensystem und Peripherie). *Naturwissenschaften* **37**: 464–476.

5

Keeping vision stable: rapid updating of spatiotopic receptive fields may cause relativistic-like effects

M. CONCETTA MORRONE, JOHN ROSS, AND DAVID C. BURR

Summary

People shift their gaze more frequently than they realize, sometimes smoothly to track objects in motion, more often abruptly with a saccade to bring a new part of the visual field under closer visual examination. Saccades are typically made three times a second throughout most of our waking life, but they are rarely noticed. Yet they are accompanied by substantial changes in visual function, most notably suppression of visual sensitivity, mislocalization of spatial position, and misjudgments of temporal duration and order of stimuli presented around the time. Here we review briefly these effects and expound a novel theory of their cause. To preserve visual stability, receptive fields undergo a fast but not instantaneous remapping at the time of saccades. If the speed of remapping approaches the physical limit of neural information transfer, it may lead to relativistic-like effects observed psychophysically, namely a compression of spatial relationships and a dilation of time.

5.1 Introduction

Saccades are ballistic movements of the eyes made to reposition our gaze. They can be deliberate but normally are automatic and go unnoticed. Not only do the actual eye movements escape notice, but so do the image motion they cause and the fact that gaze itself has been repositioned. This problem has gained the attention of most visual scientists, including von Helmholtz (1866), Sperry (1950), Alhazen (1083), and Howard (1996). A general conclusion to emerge from a variety of studies was that saccades were accompanied by a "corollary discharge" (Sperry 1950) or an "efference copy" (von Holst & Mittelstädt 1954) of the motor signal that corrected for the eye movement (for general review, see Ross et al. 2001).

There is now good evidence that many visual neurons are modified by a corollary discharge. Receptive fields of many neurons in the lateral intraparietal area (LIP) of monkeys change position before each saccadic eye movement (Duhamel et al. 1992), effectively anticipating its consequences. Similar effects have been found in other eye movement–related areas, such as superior colliculus (Walker et al. 1995), frontal eye fields (Umeno & Goldberg 1997), and even in earlier stages in the cortical visual system (Nakamura & Colby 2002), V4, V3a, and V2. The remapping of receptive fields in all these areas is consistent

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

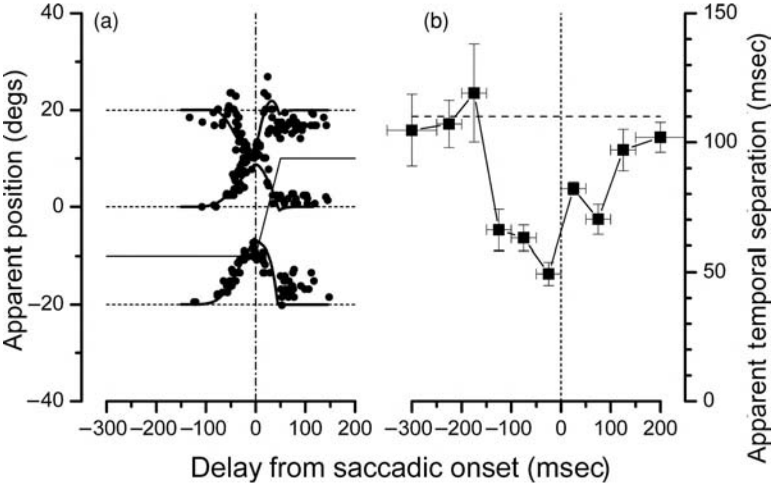


Fig. 5.1 Effect of saccades on spatial position (a) and temporal separation (b). For the spatial task, bars were briefly displayed in one of three spatial positions (−20 deg, 0 deg, and 20 deg, indicated by dashed lines). During a limited interval near the onset of the saccade, the bars were systematically misperceived toward the saccadic target, indicating compression of space at the moment of the saccade. For the temporal task, the apparent temporal separation of two briefly flashed bars was measured with a matching technique (see Fig. 5.2). At presentation times near the onset of the saccade, the apparent duration of the interval was halved, indicating temporal compression.

with psychophysical studies in the 1960s that showed that briefly displayed visual stimuli are perceived erroneously when presented around the time of saccades, displaced in the direction of the saccade (Matin & Pearce 1965; Bischof & Kramer 1968; Matin et al. 1969; Matin 1972).

5.2 Spatial and temporal mislocalization during saccades

If the errors in localization serve only to compensate for eye movements, they should always be in the same direction as the saccade. However, this is not always the case. Figure 5.1(a) (from Ross et al. 1997) shows that the direction of mislocalization of stimuli presented near the time of saccades depends strongly on the spatial position of the stimuli. Bars displayed at the center of the screen or to the left of fixation (for a rightward saccade) were displaced rightward, in the direction of the saccade. However, bars flashed beyond the saccadic target were displaced leftward, against the direction of the saccade. This results in a *compression* of visual space at the time of saccades. The compression is robust, capable of removing vernier offsets for lines flashed near saccadic onset, and causing multiple stimuli to merge into one. Compression does not occur during simulated saccades, where scenes are displaced by mirror movement with similar dynamics to saccades (Morrone et al. 1997).

More surprisingly, a recent experiment in our laboratory has shown that saccades cause a compression not only of space but also of time. Figure 5.1(b) (from Morrone et al.

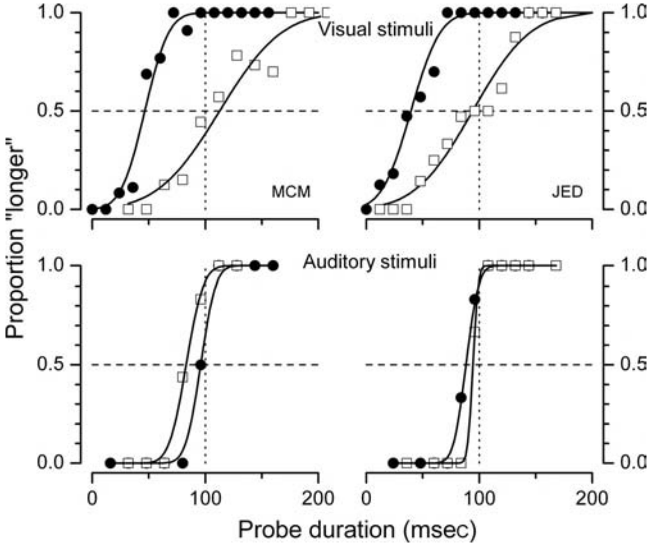


Fig. 5.2 Psychometric functions for the temporal judgments of Fig. 5.1. For the upper curves, a pair of horizontal bars separated in time by 100 msec was presented either well before (open squares) or just prior to saccadic onset (filled circles). Subjects were required to report whether this stimulus pair appeared shorter or longer than a probe pair (of variable duration) that appeared 2 sec later. With perisaccadic presentations the curves were shifted to the left, implying strong compression of time, and were steeper, implying greater precision of temporal judgments. For auditory stimuli (brief clicks) there was no temporal compression accompanying the saccade.

2005) shows how the apparent separation of two briefly flashed bars varies with time of presentation. The actual separation of the bars was 100 msec, but when flashed near saccadic onset, the apparent duration was reduced to near 50 msec. As with spatial compression, the effect was maximal near saccadic onset and followed tight temporal dynamics. The time course may appear broader than that for spatial compression, but this is only because the actual stimulus was necessarily broad (100 msec); if this is taken into account, the dynamics are similar.

Furthermore, the precision with which the judgment was made *improved* during saccades, brought out more clearly in the upper panels of Fig. 5.2 that show examples of psychophysical functions for test stimuli presented either well before saccadic onset (open squares) or perisaccadically (filled circles). Subjects were required to compare the duration of the tests (pre- or perisaccadic) with a probe pair presented 2 sec later (see Morrone et al. 2005 for full experimental details). All data are well fit by cumulative Gaussian curves. The mean of these curves estimates the *point of subjective equality* (PSE), and the standard deviations estimate the precision of the match. Note that when the stimuli are presented perisaccadically, the point of subjective equality shifts toward 50 msec, showing how time is compressed at that moment. But notice that the curves also become much

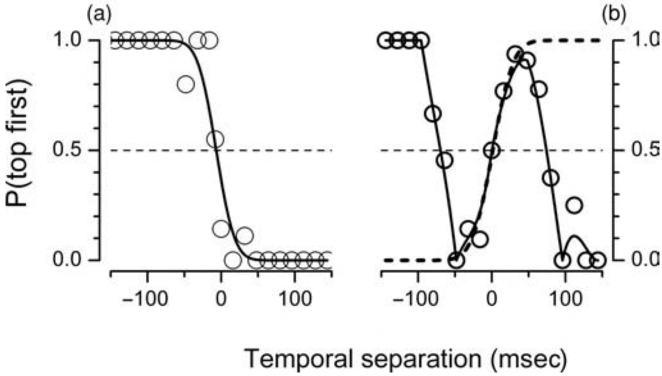


Fig. 5.3 Psychometric functions for a temporal order judgment task in which the observer was required to judge whether the top or bottom bar was presented first, as a function of temporal separation (positive times mean bottom first). The curve on the left is a standard psychometric function for bar presentations at least 100 msec after saccadic onset. The curve on the right shows data for the critical perisaccadic interval −70 to −30 msec. This psychometric function is actually inverted over the range of ±50 msec, recovering to veridicality for very large separations (the curve was fit to the circular points within the dashed lines). This implies that, for a limited range, temporal order is inverted.

steeper, showing that precision actually improves at the time of saccades. If one assumes that the precision of the judgment is determined by the neural noise associated with both test and probe stimuli, and that noise varies inversely with the number of ticks of an internal clock (Gibbon 1977), then the compression of time and the improved precision in temporal judgments are consistent with slowing of the clock during saccades.

The lower panels of Fig. 5.2 show another experiment where subjects were required to judge the duration of pairs of clicks (compared with a later probe). Here the results are quite different. Under all conditions, the judgments are more precise (in agreement with Westheimer 1999 and others), but of more relevance to this discussion, the perisaccadic and presaccadic results are indistinguishable. The effects of saccades on time are clearly modality-specific, occurring only for visual stimuli.

Even more surprisingly, for certain intervals of stimulus presentation, duration was not only misjudged but temporal order was inverted. In a further experiment, subjects were asked to estimate the temporal order of the bars (which were always presented in random order). Figure 5.3(a) shows the results for stimuli presented 100 msec or more after saccadic onset: a typical psychometric function where the probability of correct response varies smoothly with the temporal separation of the bars. However, when stimuli are presented just before saccadic onset (within the narrow range −70 to −30 msec) the psychometric function (Fig. 5.3(b)) is far from conventional but becomes triphasic. For bar separations within the range −50 to +50 msec the function runs smoothly in the opposite direction to reality – as if time had reversed. Only for large separations (greater than 100 msec or so) is temporal order perceived correctly.

5.3 Saccades and special relativity

The pairing of temporal with spatial effects that occur before saccades suggests a possible explanation: saccades induce a *relativistic* alteration of spatial and temporal metrics. In physics, relativistic effects occur when objects move at a speed approaching c, the maximal speed at which an electromagnetic wave can carry information. Propagation of information through the nervous system occurs along axons and across synapses at a speed limited by biophysical constraints. As mentioned earlier, the receptive fields of many cortical neurons are modulated by eye position (Andersen et al. 1985; Duhamel et al. 1992; Fogassi et al. 1992; Galletti et al. 1995; Duhamel et al. 1997; Nakamura & Colby 2002; Kusunoki & Goldberg 2003), presumably to anticipate and to offset the change in retinal positions. The modulation is fast but not instantaneous, often following similar dynamics to perceptual effects during saccades (Morrone et al. 1997; Ross et al. 1997; Diamond et al. 2000). As the dynamic coordinate transformation must be rapid (to offset the effects of saccades), it seems reasonable that it will occur at a rate approaching the limit of neural information transfer, and this has immediate relativistic consequences in both space and time. If these neurons mediate the perception of space and also of time (Leon & Shadlen 2003; Janssen & Shadlen 2005), transient stimuli presented during the dynamic coordinate transformation will be measured against spatial and temporal scales dilated relativistically, following the Lorentz transform (Einstein 1920): they will therefore appear compressed in one spatial dimension and in time. For neural propagations at about 87% maximum speed, objects will be compressed in both space and time by a γ factor of 2, agreeing well both with the results of spatial compression (Morrone et al. 1997; Ross et al. 1997; Lappe et al. 2000; Kaiser & Lappe 2004) and with the factor-of-two time compression reported more recently (Morrone et al. 2005). This suggestion also predicts increased temporal precision in temporal judgments during saccades, as the clock with which they are measured slows down, decreasing the number of "clock ticks" between the two events. It is also consistent with the independence on saccadic size (except for small saccades), as it is the speed of the receptive field that matters, not the duration of the shift (although the duration may affect the range over which the compression occurs). Finally, it predicts that blinks will have no effect on perceived time (although they cause many suppression-like effects that mimic saccades, with similar time course [Stevenson et al. 1986; Ridder & Tomlinson 1993]), as they are not associated with shifts in receptive fields, and that nonvisual stimuli (e.g., clicks) will not be compressed. In the following section we formally develop these ideas.

5.3.1 Maps in the brain

Images on the retina form a map (see Morgan 2003 for an interesting discussion). Let us suppose that the brain registers this map but also develops a map of activity that codes the location of objects in external space that we refer to as the craniotopic map and that this latter map is used for spatial and temporal order judgments. The craniotopic map (x', y', t') also receives a retinotopic input that will be shifted with each saccade by an amount to counteract the retinal shift in order to establish the craniotopic specificity. However, if a new signal arrives and excites this map while it is being displaced, and if the displacement takes place at high speeds, many relativistic effects of temporal and spatial localization can be predicted.

5.3.2 The Lorentz transformation

Let x, y, and t be the retinal spatial and temporal Cartesian coordinates of the stimuli (x and y can be considered signed eccentricities). Let x', y', and t' be the spatial and temporal coordinates of the craniotopic map used by our brain. The role of the craniotopic map is to assign a location of the external space and time: a stimulus that elicits a neuronal activity centered around x', y' will be localized at that external position independently of its retinal location, and the temporal separation between two peaks of activity (measured with the same neuronal clock that is used for the activity) will define the perceived delay of one stimulus over the other.

Assume that the saccade is executed horizontally from position 0 to f. Before the saccade the two maps are in register: $x' = x$, $y' = y$, and $t' = t$. After the saccade has been executed the two maps, retinal and craniotopic, will differ only by a constant spatial shift: $x' = x - f$, $y' = y$, and $t' = t$, where f is a constant that describes the position of the eye in external coordinates corresponding physiologically to the corollary discharge signal, in agreement with many models (Xing & Andersen 2000; Pouget et al. 2002). Let us suppose that, before the saccade, activity in the craniotopic map (x') begins reafferation to change its input to the retinal afference that will be appropriate to maintain craniotopic invariance after the saccade. This rapid reafference can be described mathematically as a movement in the map to reach a displacement equal and opposite to f by the end of the saccade. Let us further suppose that the motion occurs at a speed close to the maximum physiologically possible speed for the cortical area that codes the map. Maximum speed will be limited by the number of synapses involved in the transfer of information, the total length of dendrites and axons to be traveled at the diffusion speed of horizontal connection – about 0.2 m/sec (Tucker & Katz 2003) – or by a cortical-to-cortical connection of 3.5 m/sec (Girard et al. 2001).

Any signal delivered at position x, y and time t will be represented by the dynamic map at positions x', y', and t' given by the Lorentz transform

$$\begin{aligned} x' &= \gamma(x - ut) \\ t' &= \gamma(t - ux/c^2) \\ y' &= y, \end{aligned} \tag{1}$$

where

$$\gamma = \frac{1}{\sqrt{1 - u^2/c^2}} \tag{2}$$

and u is the velocity of the moving frame along the x axis (in the direction opposite to the saccade), and c is the maximum speed of neural information transmission.

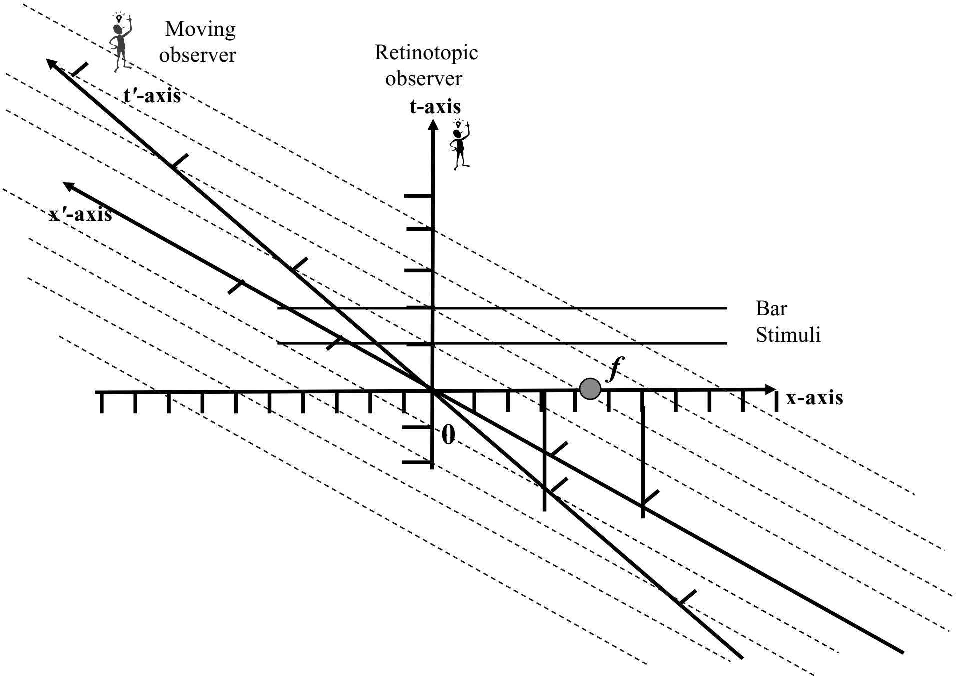


Fig. 5.4 Schematic illustration of the effect of the Lorentz transform on the stimuli of this study. Observers fixate O, then saccade to the target f that appears abruptly when the fixation spot disappears. (The saccade is assumed to occur outside the temporal limits of this graph.) At a variable time relative to saccade onset, a pair of extended horizontal bars are presented one clock tick apart. For an observer moving at relativistic speeds, in this case 0.87 c, the spatial and temporal axes (x' and t') will be rotated in space–time, following the Lorentz transformation (Equation 1). To visualize geometrically the resulting temporal dilation, consider that all lines parallel to the x' axis have constant t' (dotted lines in figure). Their intersection with the t' axis will produce a dilated time scale, with ticks more than double those of the t axis. A similar argument occurs in space, explaining spatial compression. The two horizontal gray lines represent two stimuli separated by L for the stationary map, but by L/γ for the craniotopic map.

5.3.3 Spatial compression

For simplicity, let us consider only events that take place before the eye actually moves [eye movement will alter only the actual retinal position (x) of the delivered stimuli, not the conceptual basis of the model]. Consider two brief stimuli delivered during the motion of the localization map at a spatial separation of L, depicted in Fig. 5.4 with two gray vertical lines. The distance estimate by the craniotopic map during its fast reafference will be:

$$L' = L/\gamma.$$

This expression is easy to derive if the spatial separation is evaluated at the same temporal instant t' of the craniotopic activity. (To derive this expression just calculate Equation 1 for difference of two impulses and assign at $\Delta x = L$ and $\Delta t' = 0$).

If we consider that $c = 1$ and $u = 0.87c$, γ will be equal to 2, so

$$L' = L/2.$$

This implies that apparent distance would appear compressed by about a factor of two (agreeing well with the psychophysics: Morrone et al. 1997; Ross et al. 1997; Lappe et al. 2000; Kaiser & Lappe 2004). If only one stimulus is delivered, its position may be judged by the distance between the stimulus and one of the activities present in the dynamic map, usually the saccadic target, obtaining the same compression for stimulus location as for separation. If the speed of the dynamic remapping is low compared to maximum speed and equal to $\Delta f/\Delta t$, where Δt is the typical saccadic mislocalization temporal dynamics and Δf the cortical distance covered by the saccade, the distance of the two targets will be unaffected and their positions will be shifted in the direction of the saccade, as sometimes observed in the dark for memory-guided saccades (Cai et al. 1997; Lappe et al. 2000).

The major spatial compression takes place within a 50-msec interval around saccadic onset (for a saccade of 20 deg). This implies a speed of information transfer of 400 deg/sec and an estimate of the maximum speed of about 440 deg/sec.

5.3.4 Compression of time

A similar argument can be developed for the temporal judgments. The stimuli presented in this experiment encompass all spatial positions along x, although the two bars will excite different pools of neurons given their vertical offset. Each bar will be interpreted by the dynamic moving map as a stimulus encoded at different spatial positions (x') and time (t'). If we assume (conservatively) that the temporal judgments are performed by neurons that assume the same x' position of the moving map, the resulting estimate will be

$$T' = T/\gamma = 100\,\text{msec}/2 = 50\,\text{msec},$$

where T and T' are the measured temporal separations of the stimulus pair in the two different inertial frames of reference. (Again, to obtain this expression it is sufficient to repeat the procedure used for the spatial compression and assign $\Delta x' = 0$ and $\Delta t = T$).

Given that each tick of the clock of the moving map lasts two ticks of the stationary (retinally anchored) clock, we should also have a decreased error if error is associated with number of ticks to be counted (Weber's law). This estimate holds when both stimuli are delivered during the dynamic phase of the map. In our condition this corresponds to when one bar is delivered 50 msec before the eye movement and the other 50 msec after the eye movement. To simulate the interval before or well after saccadic onset, we could consider that only some of the clock ticks are subject to time dilation induced by the dynamic remapping and that this would necessarily decrease the time compression effect.

At some crucial intervals just before saccadic onset, the temporal order of the bars was consistently seen to be inverted. The reversal of apparent order does not follow immediately from special relativity alone but could easily be accounted for within this framework by considering temporal *postdiction* (Libet et al. 1979; Eagleman & Sejnowski 2000). If each of the two successive bars is labeled independently and referred backward in time by *N* clock ticks to compensate for delays in neural processing, the second bar will sometimes be pushed backward beyond the first if the clock were ticking more slowly at the time of second bar presentation.

5.4 Concluding remarks

The idea of a shift of reference to preserve visual stability is an old one: what is novel here is the suggestion that a shift can and does have relativistic consequences. This suggestion can explain why spatial compression is predominantly one-dimensional and why it is accompanied by temporal compression and reversal of temporal order. Shifts of reference are not rare: they occur whenever saccades are made, three times a second on average throughout waking life. The effects of saccades, common though they are, escape notice because they are nullified by an accompanying shift of frame of reference. Only transient stimuli, rare in nature, will be distorted, and then only if they happen to appear while the rapid shifts of reference are in progress.

Acknowledgment

This research was supported by the European Union Framework Programme 6 and 7: “MEMORY” and ERC advanced grant 229445 “STANIB.”

References

Alhazen, I. (1083). Book of optics. In A. I. Sabra (ed.), *The Optics of Ibn al-Haytham.* London: Warburg Institute.

Andersen, R. A., Essick, G. K., & Siegel, R. M. (1985). Encoding of spatial location by posterior parietal neurons. *Science* **230**(4724): 456–458.

Bischof, N., & Kramer, E. (1968). Untersuchungen und Überlegungen zur Richtungswahrnehmung bei wilkuerlichen sakkadischen Augenbewegungen. *Psychol Forsch* **32**: 185–218.

Cai, R. H., Pouget, A., Schlag-Rey, M., & Schlag, J. (1997). Perceived geometrical relationships affected by eye-movement signals. *Nature* **386**: 601–604.

Diamond, M. R., Ross, J., & Morrone, M. C. (2000). Extraretinal control of saccadic suppression. *J Neurosci* **20**: 3442–3448.

Duhamel, J., Bremmer, F., BenHamed, S., & Graf, W. (1997). Spatial invariance of visual receptive fields in parietal cortex neurons. *Nature* **389**: 845–848.

Duhamel, J. R., Colby, C. L., & Goldberg, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science* **255**(5040): 90–92.

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science* **287**(5460): 2036–2038.

Einstein, A. (1920). *Relativity: The Special and General Theory.* New York: Henry Holt.

Fogassi, L., Gallese, V., di Pellegrino, G., Fadiga, L., Gentilucci, M., Luppino, G., et al. (1992). Space coding by premotor cortex. *Exp Brain Res* **89**(3): 686–690.

Galletti, C., Battaglini, P. P., & Fattori, P. (1995). Eye position influence on the parieto-occipital area PO (V6) of the macaque monkey. *European J Neurosci* **7**: 2486–2501.

Gibbon, J. (1977). Scalar expectancy theory and Weber's Law in animal timing. *Psychol Rev* **84**: 279–325.

Girard, P., Hupe, J. M., & Bullier, J. (2001). Feedforward and feedback connections between areas V1 and V2 of the monkey have similar rapid conduction velocities. *J Neurophysiol* **85**(3): 1328–1331.

Howard, I. P. (1996). Alhazen's neglected discoveries of visual phenomena. *Perception* **25**(10): 1203–1217.

Janssen, P., & Shadlen, M. N. (2005). A representation of the hazard rate of elapsed time in macaque area LIP. *Nat Neurosci* **8**(2): 234–241.

Kaiser, M., & Lappe, M. (2004). Perisaccadic mislocalization orthogonal to saccade direction. *Neuron* **41**(2): 293–300.

Kusunoki, M., & Goldberg, M. E. (2003). The time course of perisaccadic receptive field shifts in the lateral intraparietal area of the monkey. *J Neurophysiol* **89**(3): 1519–1527.

Lappe, M., Awater, H., & Krekelberg, B. (2000). Postsaccadic visual references generate presaccadic compression of space. *Nature* **403**: 892–895.

Leon, M. I., & Shadlen, M. N. (2003). Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron* **38**(2): 317–327.

Libet, B., Wright, E. W., Jr., Feinstein, B., & Pearl, D. K. (1979). Subjective referral of the timing for a conscious sensory experience: a functional role for the somatosensory specific projection system in man. *Brain* **102**(1): 193–224.

Matin, L. (1972). Eye movements and perceived visual direction. In D. Jameson & L. M. Hurvich (eds.), *Handbook of Sensory Physiology vol. VII/Visual Psychophysics* (331–380). Berlin: Springer-Verlag.

Matin, L., Matin, E., & Pearce, D. G. (1969). Visual perception of direction when voluntary saccades occur: I. Relation of visual direction of a fixation target extinguished before a saccade to a subsequent test flash presented during the saccade. *Perception and Psychophysics* **5**: 65–68.

Matin, L., & Pearce, D. G. (1965). Visual perception of direction for stimuli flashed during voluntary saccadic eye movements. *Science* **148**: 1485–1487.

Morgan, M. J. (2003). *The Space between Your Ears: How the Brain Represents Visual Space*. London: Weidenfeld & Nicolson.

Morrone, M. C., Ross, J., & Burr, D. C. (1997). Apparent position of visual targets during real and simulated saccadic eye movements. *J Neurosci* **17**: 7941–7953.

Morrone, M. C., Ross, J., & Burr, D. (2005). Saccadic eye movements cause compression of time as well as space. *Nat Neurosci* **8**(7): 950–954.

Nakamura, K., & Colby, C. L. (2002). Updating of the visual representation in monkey striate and extrastriate cortex during saccades. *Proc Natl Acad Sci U S A* **99**(6): 4026–4031.

Pouget, A., Deneve, S., & Duhamel, J. R. (2002). A computational perspective on the neural basis of multisensory spatial representations. *Nat Rev Neurosci* **3**(9): 741–747.

Ridder, W. H. III, & Tomlinson, A. (1993). Suppression of contrast sensitivity during eyelid blinks. *Vision Res* **33**(13): 1795–1802.

Ross, J., Morrone, M. C., & Burr, D. C. (1997). Compression of visual space before saccades. *Nature* **384**: 598–601.

Ross, J., Morrone, M. C., Goldberg, M. E., & Burr, D. C. (2001). Changes in visual perception at the time of saccades. *Trends Neurosci* **24**: 113–121.

Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *J Comp Physiological Psych* **43**: 482–489.

Stevenson, S. B., Volkmann, F. C., Kelly, J. P., & Riggs, L. A. (1986). Dependence of visual suppression on the amplitudes of saccades and blinks. *Vision Res* **26**(11): 1815–1824.

Tucker, T. R., & Katz, L. C. (2003). Spatiotemporal patterns of excitation and inhibition evoked by the horizontal network in layer 2/3 of ferret visual cortex. *J Neurophysiol* **89**(1): 488–500.

Umeno, M., & Goldberg, M. (1997). Spatial processing in the monkey frontal eye field. I. Predictive visual responses. *J Neurophysiol* **78**: 1373–1383.

von Helmholtz, H. (1866). Handbuch der physiologischen optik. In J. P. C. Southall (ed.), *A Treatise on Physiological Optics*. New York: Dover.

von Holst, E., & Mittelstädt, H. (1954). Das reafferenzprinzip. *Naturwissenschaften* **37**: 464–476.

Walker, M. F., Fitzgibbon, J., & Goldberg, M. E. (1995). Neurons of the monkey superior colliculus predict the visual result of impending saccadic eye movements. *J Neurophysiol* **73**: 1988–2003.

Westheimer, G. (1999). Discrimination of short time intervals by the human observer. *Exp Brain Res* **129**(1): 121–126.

Xing, J., & Andersen, R. A. (2000). Models of the posterior parietal cortex which perform multimodal integration and represent space in several coordinate frames. *J Cogn Neurosci* **12**: 601–614.

6

Combined influences of extraretinal signals, retinal signals, and visual induction on space perception and manual behavior in perisaccadic and steady viewing

LEONARD MATIN AND WENXUN LI

Summary

The locations of stationary objects appear invariant, although saccadic eye movements shift the images of physically stationary objects on the retina. Two features of this perceptual stability related to saccades are that postsaccadic locations of objects appear invariant relative to their appearance in the presaccadic view, and perception of postsaccadic stimulation is free from interference by remnants of presaccadic stimulation. To generate stability, quantitatively accurate cancellation between retinal input (RI) and extraretinal eye position information (EEPI) must occur, and persisting influences from the presaccadic view must be eliminated. We describe experiments with briefly flashed visual stimuli that have measured (1) the time course of perisaccadic spatial localization, (2) the interfering effects of persisting stimulation prior to the postsaccadic period, (3) the achievement of perceptual stability by removing visual persistence early, and (4) the influence of metacontrast utilizing the normal perisaccadic spatiotemporal distribution of retinal input to prevent interference from visual persistence.

For the steady eye, a generalized cancellation mechanism is analyzed through studying mislocalizations in perceptual orientation and visually guided manual behavior produced by (1) modifying EEPI in observers with experimental partial paralysis (curare) of the extraocular muscles and/or (2) modifying RI by varying visual field orientation (i.e., its pitch and/or roll). The influences of visual pitch and roll derive from the retinal orientations of individual straight lines and their combinations, with the identical lines influencing perceived verticality and elevation. Surprisingly, although full accuracy is found in pointing and orienting a fully extended finger/hand/arm to perceptually mislocalized elevations and frontal-plane orientations, large errors in manual settings equal and opposite to the perceptual errors occur with the hand close to the body. This variation in manual accuracy is linearly graded with hand-to-body distance and serves useful ecological purposes at both extremes as well as at the intermediate distances.

6.1 Introduction

None of the visual system's extraordinarily versatile capabilities is more remarkable than the ability to perceptually localize objects in egocentric space and to carry out accurate, visually guided sensorimotor behavior in reaching for them. Most significant for our adaptation to

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

life on earth is the ability to maintain perceptual constancy of an object's spatial location in the face of movements of the eye while effortlessly distinguishing movements of objects in the environment from movements by the observer's own eye that generate identical stimulations to the retina. However, generally unnoticed, but of equally great significance for perceptual constancy of spatial location and orientation as well as manually guided sensorimotor behavior, is the fact that such perceptual constancy is normally achieved in a visual environment whose orientation is consistent with the direction of gravity, providing a powerful frame of reference for perception and for action.

This article describes the mechanism that maintains perceptual constancy when saccades occur in an erect visual environment and then goes on to describe the maintenance of perceptual constancy based on the consistency between orientation of the visual framework and the direction of gravity.

6.2 Spatial localization and saccadic eye movements

Much of the focus on spatial constancy in the presence of eye movements has been on the contribution of extraretinal processes to shifting Lotze-like retinal local signs (1886) for visual direction, with emphasis on whether these processes are fed forward from the command system controlling movements of the eye (outflow signal) or fed back from peripheral organs in the orbit (inflow signal) (see Matin 1972, 1976a, 1982, 1986; Shebilske 1977; Li & Matin 1997; Schlag & Schlag-Rey 2002, for reviews). But the pre-1960 treatment did not include the complex involvement of processes originating in the visual nervous system that play a significant role in maintaining perceptual stability in the presence of saccades.

6.2.1 Uncovering EEPI with flashes of light

The visual system did not evolve to cope with brief flashes of small isolated points of light in otherwise total darkness. However, such stimulation has been central to the methodology in investigations relating perceptual stability and eye movements as the most straightforward way for experimenters to localize stimulation in both time and space while maintaining reasonably constant adaptation without complicating stimulation from a steady background, and because considerable perceptual stability had already been demonstrated in similar experiments during attempts at maintaining steady fixation in total darkness (Matin & Kibler 1966; Matin et al. 1966). The use of such stimulation during saccades forced recognition of the care needed in dealing with three additional matters relating to the retinal side of things: (1) visual persistence, (2) retinal smears, and (3) visual latencies.

In the first experiment directly aimed at demonstrating the existence of an extraretinal signal regarding eye position (i.e., EEPI: extraretinal eye position information) that would stabilize spatial localization in the presence of a voluntary saccade (Matin & Pearce 1965), a brief test flash was presented during a saccade, and the subject compared its location to a target viewed before the saccade (Fig. 6.1). If the flash was perceived at the same physical location as the earlier comparison target, it could be interpreted as demonstrating the presence of a temporally sampled version of the normal continuous view, perceptually stabilized by the time-varying extraretinal signal. In the first of two conditions, the subject compared the location of the test flash to that of the saccadic target (upper graph); in the second condition the subject compared the location of the test flash to that of the fixation target (lower graph). To eliminate influences from the visual field, total darkness was employed (invisible infrared beams were employed in the highly sensitive and accurate contact-lens eye-movement monitor; Matin 1964; Matin & Pearce 1964) and as much time separated the three necessary stimuli (Fig. 6.1(c)) as the work permitted. The limitations on time resulted from concerns about the possibility of too much involuntary eye movement in total darkness before the saccade if the fixation target was gone too long, and too much visual persistence if stimuli followed each other too quickly. Instead of results disclosing the extraretinal signal, in the first condition the point of subjective equality (PSE) for the saccadic target was close to a retinotopic match (Fig. 6.1(a)). However, a comparison of the location of the test flash to the previously viewed fixation target (Fig. 6.1(b)) yielded the originally anticipated result. In order for the test flash that was presented in the middle of the saccade to appear at the same location as the previously viewed fixation target, the test flash had to be nearly at the physical location of the fixation target itself. This provided support for a stabilizing extraretinal signal in an experimental context in which the process had been interrogated without completely destroying it.

6.2.2 The sluggish change in perisaccadic spatial localization

Two matters became clear in subsequent experiments (Matin & Matin 1964–1965, unpublished) aimed at disentangling the influences of the temporal and location values of the three stimuli that were fixed parameters in the two conditions in Fig. 6.1: (1) The difference in results for the two comparison targets in Fig. 6.1 was not a consequence of the target's function as a fixation target or a saccadic target (also see Section 6.2.7). (2) Changes in the temporal parameters of the visual stimulation alone could influence the results drastically.

To delineate these temporal parameters in an expansion of the paradigm (see Fig. 6.1(c)), the test flash was presented at various times before, during, or following the 2° 11' saccades. Figure 6.2 displays the average measurements on three subjects, showing the time course of EEPI's development as measured by the shift at the retina (retinal PSE) required for a test flash to appear in the same location as the fixation target (Matin et al. 1969, 1970; Matin & Matin 1972; Matin 1976b). The 1-msec variably-located test flash was presented each of a number of times from about 300 msec prior to the saccade to about 1500 msec following the saccade. Continuous measurements of horizontal eye movements were made throughout each trial. The retinal PSE (ordinate) changed slowly relative to the saccade itself throughout the range of durations. Thus, the existence of a contribution of EEPI originally found (Fig. 6.1) was confirmed. But these experiments failed to substantiate a model consisting of visual and extraretinal components that were assumed to maintain equal and opposite

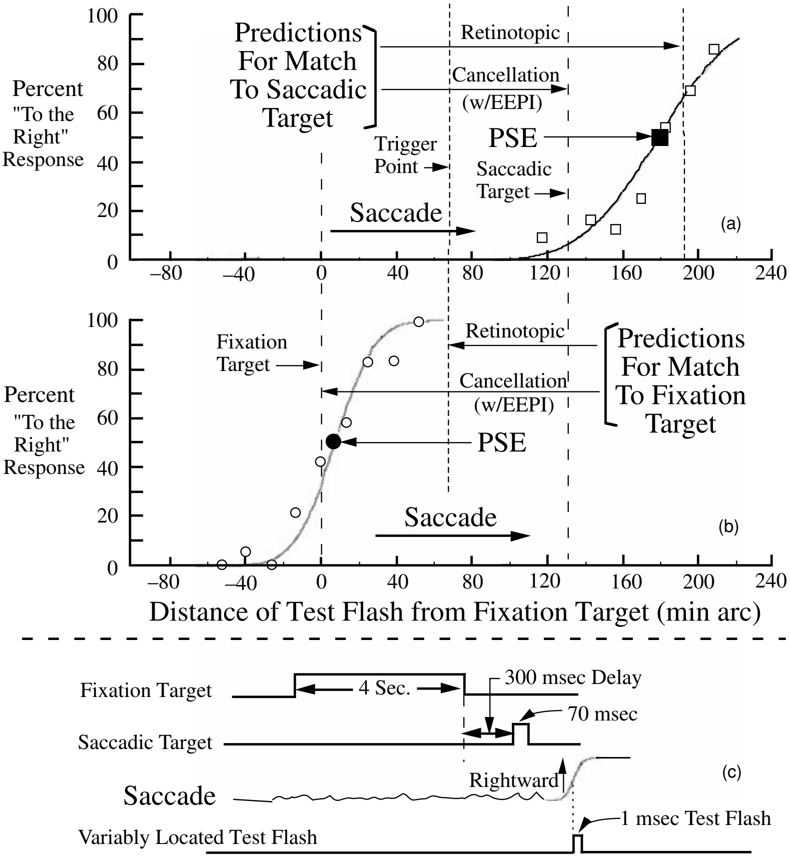


Fig. 6.1 Summary of the psychophysical results in two experiments (Matin & Pearce 1965) in which the subject made voluntary horizontal 2°11′ saccades in total darkness employing the paradigm in (c). The variably located test flash (abscissa) was presented when the fovea crossed the trigger point halfway between the fixation and saccadic targets. The flash was triggered by an electronic signal from the infrared contact-lens apparatus continuously measuring eye position (Matin 1964; Matin & Pearce 1964). In (a) the subject reported the location of the test flash relative to the location of the saccadic target; in (b) the subject reported the location of the test flash relative to the location of the fixation target. The ordinate values of the open squares in (a) and open circles in (b) display the percentage of trials on which the test flash appeared to the right of the comparison target. PSEs (points of subjective equality) were calculated from the best-fitting normal given in each case. When the subject's comparison was to the saccadic target (a), the PSE was close to a retinotopic match. When the comparison was to the fixation target (b), the PSE was close to an accurate match. The cancellation (w/EEPI) locus shown for each condition indicates where full perceptual accuracy would be (see text); the retinotopic locus is the target location for which the test flash would have stimulated the same retinal locus as the comparison target.

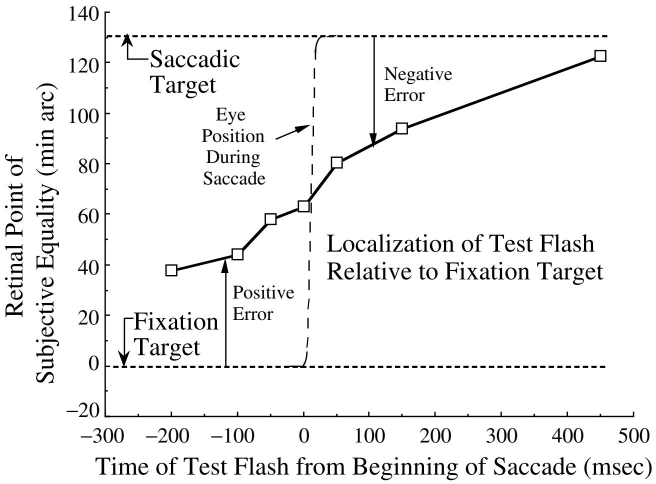


Fig. 6.2 The paradigm in Fig. 6.1(c) was employed to trigger a test flash at various times before, during, or after the saccade. The subject compared the horizontal location of the test flash to that of the fixation target in a method of constant stimuli format. PSEs (points of subjective equality) were calculated from the psychometric function at each time. The values displayed are the averages from three subjects. (Data from Matin et al. 1969, 1970; Matin 1972, 1976, 1986; Matin & Matin 1972.) Labeling the errors in the saccade direction (prior to the saccade) as "positive" and errors in the opposite direction (following the saccade) as "negative" in this figure follows the usage introduced by Honda (1993) that has become standard in some quarters (see Section 6.2.6).

magnitudes in the temporal neighborhood of the saccade ("cancellation model") and would require a constant difference between them to attain a constant localization value. If EEPI had tracked the saccade and was responsible for perceptual stability, it should have been possible to interrupt the visual stimulus, eliminate the portion presented during the saccade, and obtain stability for a postsaccadic stimulus as well as for a presaccadic stimulus relative to perception prior to the saccade. But, as the results in Fig. 6.2 show, neither presaccadic nor postsaccadic stability was obtained. Subsequent experiments with 5-deg-long and 8-deg-long saccades were in agreement with the monotonic shift in Fig. 6.2 (Pola 1973, 1976). The normal stimulation contained in the saccadic stimulus contributed something essential to stability, and elimination of it prevented stability (see Section 6.2.5). The trial-to-trial variability in oculomotor behavior in these experiments led to the discovery of another aspect of the control by EEPI: for any given time following the saccade, the retinal PSE for the previously viewed fixation target was systematically dependent on the position of the eye.

Both Hallett and Lightstone (1976a,b) and Hansen and Skavenski (1985) later reported high levels of accuracy in making eye movements and arm movements to targets during voluntary saccades, thus raising questions about how motor behavior could be so accurate if perceptual localization manifested such sizeable errors as in Figs. 6.1 and 6.2. However, subsequent work from three independent laboratories failed to substantiate the high levels of motor accuracy in pointing with the eye or the hand, but instead, in experiments involving perceptual localization, manual localization, and direction of gaze, motor errors manifested a similar sluggishness to perceptual errors, and the sluggishness in both was similar to that seen in the earlier experiments with perceptual localization alone (Honda 1989, 1990, 1991, 1993, 1999; Dassonville et al. 1992, 1993, 1995; Miller 1996; Bockisch & Miller 1999; Schlag & Schlag-Rey 2002).

6.2.3 Constraints on EEPI from visual persistence

The initial comparison between the perceived locations of the test flash and the fixation target had produced evidence for an extraretinal signal regarding eye position, but the comparison to the saccadic target had not (Fig. 6.1). It was conjectured (Matin & Pearce 1965; Matin 1972) that the difference was a consequence of (a) a difference in visual persistence between the fixation target and the saccadic target when the comparisons to the test flash were made, and (b) the persistence difference that arose from the difference in temporal interval between the fixation-target-offset/test-flash-onset (approximately 500 msec) and the saccadic-target-offset/test-flash-onset (approximately 100 msec). The much briefer saccadic-target/test-flash interval led to a cotemporal period for persistence of the two items in the comparison that resulted in retinotopy, whereas the lack of such a cotemporal period for the fixation target permitted EEPI to work. Following experimental confirmation of the conjecture (Fig. 6.3) (Matin et al. 1971; Matin 1972, 1976b), a generalization (Matin 1972, p. 353) was proposed:

> The influence of a saccade-contingent extraretinal signal on relative visual direction of different stimuli is severely restricted unless a sufficiently long time interval separates their presentations. Relative visual directions of successive stimulations presented at shorter time intervals appear to be essentially judged on the basis of relative retinal location alone as in the steadily fixating eye. The short interval effects evidently depend on whether the neural consequences of the earlier stimulation can persist into the time period during which the second stimulus is being processed. This persistence will undoubtedly depend on variables such as state of light adaptation and stimulus intensity . . . we may conceive of two storage ‘registers’: a ‘short-term’ memory storage and a ‘long-term’ storage. When a saccade occurs, the extraretinal signal associated with it may influence the relation of visual directions of two targets only if the neural process corresponding to [the first] one of them exists in the long-term storage when the [second] one is seen.

In the first part of a two-part experiment (Fig. 6.3), the during-the-saccade and postsaccadic segments of the earlier paradigm were repeated (one-flash paradigm, Fig. 6.3(b)); the variably located 1-msec test flash was used to measure the retinal PSE for the previously viewed fixation target following different postsaccadic delays (unfilled circles, Fig. 6.3(a)). In the second part (two-flash paradigm, Fig. 6.3(c)) the 0-delay retinal PSE from the first part was presented when the eye crossed the trigger point, and served as

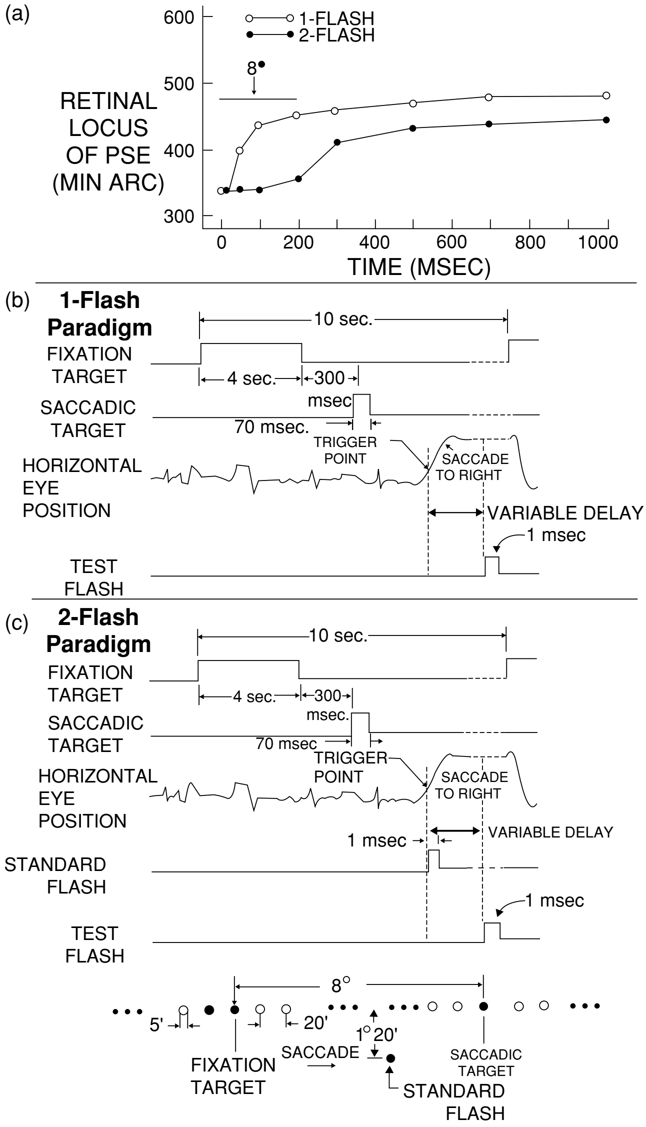


Fig. 6.3 (a) Results for an experiment in which the influence of visual persistence on localization of a target flashed in conjunction with 8-deg-long saccades is systematically varied. (b) Sketches the *one-flash paradigm* in the first part of the experiment. (c) Sketches the *two-flash paradigm* in the second part of the experiment. The one-flash (open circle) data displayed in (a) are measurements of the retinal PSE for the previously viewed fixation target as measured by the variably located test flash using the paradigm in (b). The two-flash (filled-circle) data displayed in (a) are the retinal PSEs for the standard flash as measured by the variably located test flash using the paradigm in (c). *In the second part the standard flash was presented as the eye crossed the trigger point; it was presented 1°20' below the horizontal location that had been determined as the PSE in the first part of the experiment when the eye had been at the trigger point (0-delay).* Note the early approximate 200-msec period during which the retinal PSE for the standard flash (*one-flash paradigm* in the first part of the experiment) is essentially constant and thus retinotopic, demonstrating the presence of persistence; also note the subsequent gradual PSE rise as visual persistence of the first flash dissipates in the two-flash paradigm in the second part of the experiment (Matin et al. 1971; Matin 1976b).

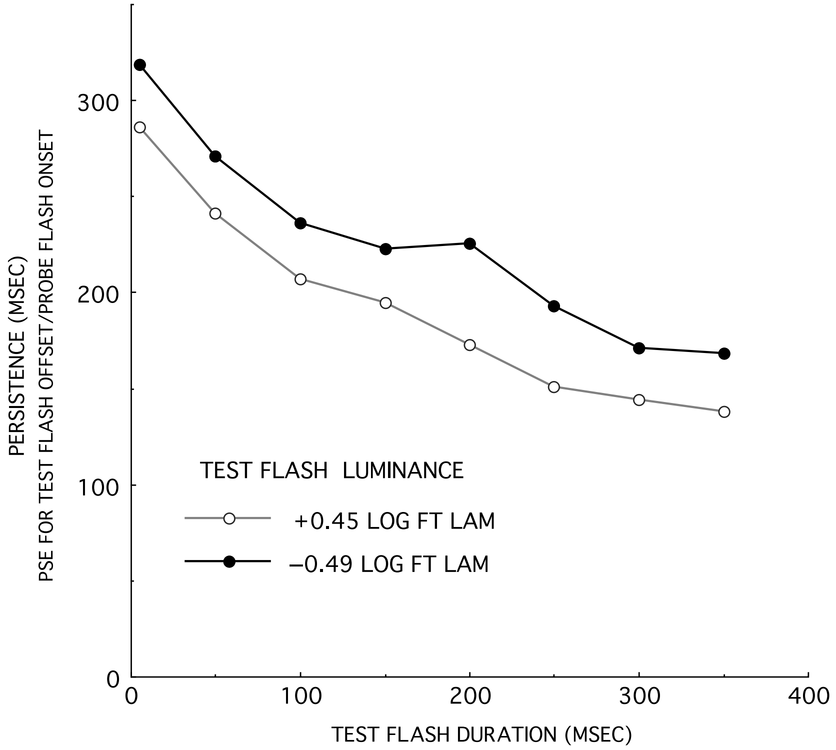


Fig. 6.4 Visual persistence (PSE for test flash offset/probe flash onset) as a function of test flash duration at each of two luminance levels (average for two Ss; Bowen et al. 1974).

the standard flash against which the location of the test flash – presented at a variable delay from the standard flash – was judged. This match with the standard flash in the second part (filled circles, Fig. 6.3(a)) was essentially constant and retinotopic until the test flash was presented about 200 msec following the standard flash, at which time the retinal PSE began to change along a time course that was somewhat slower than that in the first part. This 200-msec retinotopic period corresponds to the duration that visual persistence of the standard flash overlaps the period during which the test flash was visible; the subsequent growth along a slower time course is consistent with a gradual decrease in the persistence of the standard flash beginning near the end of the 200-msec period.

Direct psychophysical measurements of visual persistence with the stationary eye demonstrated the presence of a significant second component to the persistence explanation for the difference between the two results in Fig. 6.1. An increase in visual persistence with decrease in flash duration as measured for the flash by offset/onset PSEs relative to a constant probe flash is seen in Fig. 6.4 (Bowen et al. 1974; see also Efron 1970; Bowen 1981). The longer persistence for the short-duration (70 msec) saccadic target over the long-duration (4 sec) fixation target results in a difference in persistence approximating 100 msec and adds to the influence of the difference in temporal interval; this further strengthens the persistence explanation of the nearretinotopic result in Fig. 6.1. Systematic decreases in persistence have been measured with increases in adapting luminance of a background (Matin & Bowen 1976; Bowen 1981), but persistence against high-level backgrounds remains substantial, indicating the significance of the persistence interpretation in visible backgrounds as well as in darkness.

6.2.4 Further persistence-related work

Additional support for the persistence explanation derives from work in four laboratories. Experiments made use of brief flashes during trials in which voluntary saccades were made; all involved manipulations of the temporal parameters of the stimulation, and two of them employed differences in luminous intensities between different segments of the stimulation during an individual trial. Although it is clear that significant variations in visual latency were present in all four sets, it is not clear that these latency variations played the role that might have been expected. However, all four sets are readily explained by the contributions of visual persistence and its timing.

a. Flashed Vernier offsets and visual latencies (excerpts by Matin et al.; described in Matin 1976b)**:** An experiment based on one of the better-known facts of visual science – the decrease in response latency with increased intensity of visual stimulation – called for vernier discriminations between two vertical lines of identical or widely disparate intensities presented simultaneously during a saccade. Subjects executed a $2^\circ 11'$ horizontal saccade during which two $34'20'' \times 1'5''$ vertical lines vertically separated by $2'20''$ were flashed simultaneously for 1 msec at the moment that the fovea crossed the trigger point midway between the fixation and saccadic targets. In two of the four conditions, the intensity of the two lines differed by a factor of 100, with the intensity of the upper line higher in one condition and lower in the other; in the other two conditions the two lines were presented at the same high intensity or the same low intensity. The magnitude and direction of horizontal offset between the two lines was randomly varied among trials in each condition, and the subject reported on the horizontal offset between the two lines. No changes were made between conditions other than the insertion, removal, or shifting of two-log unit neutral density filters between filter holders in the upper and lower beams.

The total extent of each of the four psychometric functions for each subject was less than 1 minarc with no discernible systematic differences among the four functions. Our calculations from other work indicate that, at the intensities involved, the 2-log unit difference would produce at least a 10- to 30-msec difference in onset latency. The saccade's duration was about 25 msec (from 10% to 90% complete), with a velocity greater than 6 minarc/msec in the middle of the saccade when the two lines were simultaneously flashed. Thus, the eye would have moved at least 1 deg in 10 msec, and if EEPI had kept pace with the eye, the subjects should have seen the two collinear lines displaced by at least 1 deg. This is a far cry from the measured nonsystematic less-than-1-minarc total variation of the four psychometric functions. Where did the onset latency difference go? The result strongly suggests that the long cotemporal segments of the persisting neural ("retinal") responses of the two brief flashes forced retinotopic comparison related to the cotemporal period (as in the 0- to 200-msec "retinotopic period" in Fig. 6.3(a), section 6.2.3), preventing EEPI from operating separately on the two lines.

b. Presaccadic flash duration and 3-point Vernier alignment (Cai et al. 1997; also see Miller & Bockisch 1997): When long-duration and short-duration flashes to different retinal loci terminated simultaneously shortly before a saccade in a 3-point vertical vernier task, the short-duration center dot appeared horizontally displaced from the longer-duration two outer dots in the direction of the impending saccade by about 1 deg, although no such displacement occurred when a saccade was not impending. The explanation for this interesting result follows immediately from the combination of two of the facts described above: (a) visual persistence is considerably longer for short-duration flashes than for long-duration flashes (Bowen et al. 1974, Fig. 6.4); (b) the magnitude of mislocalization for stimulation in the presaccadic period grows monotonically with time in the 200 msec before a saccade, reaching a presaccadic maximum at saccade onset (Matin et al. 1970; Matin 1972; Fig. 6.2(a)). Because visual persistence for the long-duration flash is over and gone considerably before that of the short-duration flash in the presaccadic period, the growing EEPI related to the impending saccade coexists with the longer-persisting short flash for a longer time period than it does with the shorter-persisting long flash, and thus the EEPI for the short flash is of greater magnitude than for the long flash, resulting in a larger shift of perceived location for the short flash.

c. Luminance insensitivity (Boucher et al. 2001): Consistent with the insensitivity to the visual latency/intensity variation in the two-flash vernier experiment (Section 6.2.4a), differences in perisaccadic mislocalization due to luminance differences are much smaller than models based on differences in visual latency can account for. The authors obtain some improvements in prediction by assuming large magnitudes of damping of EEPI, thus reducing the theoretical influences of latency differences, although they consider these magnitudes implausible. It is likely that such damping may be the expression of increased visual persistence consistent with the approach we present here.

d. Effects of interflash interval (Sogo & Osaka 2001, 2002): In the first of the authors' extensive series of experiments, retinotopic localization was found with a short interflash interval (78 msec) and nonretinotopic mislocalization with a long interval (2 sec) (Sogo & Osaka 2001). They suggest a basis in differences in perceptual strategy that is exocentric with short intervals and egocentric with long ones, where the egocentric strategy involves EEPI and the exocentric strategy does not. In 2002, they reported an extension of their retinotopic upper limit to 120 msec and reduction of the egocentric lower limit to 240 msec. Given the differences in parameters and method, their results are entirely consistent with those in Fig. 6.3, and their differences between short and long interflash intervals are consistent with the visual persistence differences described above.

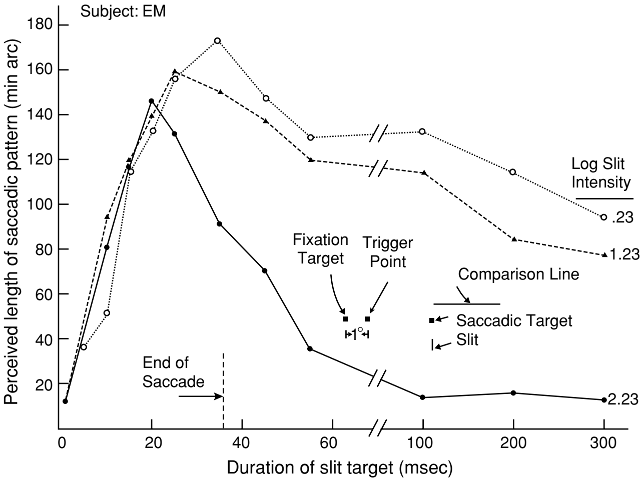


Fig. 6.5 The perceived length of the saccadic stimulus of a transiently transilluminated slit (2' horizontal × 30' vertical) is displayed as a function of the duration of slit illumination. Viewing was monocular against a completely dark field except for the fixation and saccadic targets above the slit. The onset of the slit was triggered from the eye movement monitor at the moment that the eye crossed the trigger 1 deg into the 4 deg saccade and remained illuminated for the duration shown on the abscissa. The comparison line was presented 350 msec following the saccade for 2 msec (method of constant stimuli) to obtain the perceived smear length (Matin et al. 1972).

6.2.5 Saccadic suppression and metacontrast

If the presaccadic view stimulating the retina were to persist into the saccadic period and beyond, it would interfere with clear vision of the postsaccadic view, much like two successive exposures of the same segment of film in a camera that had been rotated between the exposures; the problems are similar for eye and camera exposures during rotation. Thus, it is essential to eliminate visual persistence prior to the postsaccadic period if vision is to possess its normal clarity and for EEPI to do the important work of stabilizing spatial localization in the postsaccadic period. In fact, the clearing mechanism is effective, and interference from such double exposures and saccadic smears only occurs in rare and/or unusual circumstances (E. Matin 1974). The main mechanism for clearing visual persistence of earlier stimulation is a consequence of the spatiotemporal course of retinal stimulation near a saccade and is carried out by metacontrast, a mechanism that had been measured with the stationary eye and its pervasiveness demonstrated in the somesthetic and auditory systems as well as in vision (Alpern 1952, 1953; Kolers & Rosner 1960; Kolers 1962; Raab 1963; von Békésy 1967, 1968; Kahneman 1968; Weisstein 1972; Breitmeyer & Ganz 1976). This was discovered in experiments (Fig. 6.5) in which exposure of a stationary narrow vertical slit began with the eye 1 deg into a 4-deg rightward-going horizontal saccade in a dark field, and was terminated following different exposure durations with the eye either further into the saccade or following its completion (E. Matin et al. 1972). The perceived horizontal extent of the slit was measured by comparison with a 2-msec variable-length horizontal line presented to the stationary eye 350 msec following the saccade. Perceived length grew with exposure duration, matching the retinal extent stimulated by the slit on the saccading eye up to exposure durations close to the saccadic duration. However, for longer exposures perceived length diminished until, for sufficiently long exposures, it again appeared as a narrow slit as it did for very brief exposures during the saccade or for continuous viewing with the stationary (nonsaccading) eye. As a segment of their subsequent work containing other interesting observations, Campbell and Wurtz (1978) replicated these results.

Although the magnitude of illumination by the fixed-intensity light is distributed across the retina inversely with saccade velocity (peak velocity is near the center of the saccade), following saccade termination, the energy in a continuing exposure accumulates on a single retinal region. This accumulation provides the metacontrast induction stimulus generating suppression and shortening of perceived length as this later, more intense segment inhibits visibility in the earlier, less intense segment ("backward masking") whose persistence has also begun to diminish on its own. The growth of smear length with exposure duration up to the 35-msec duration of the saccade matched the length of the retinal distribution, but the length peak was earlier and shorter with the more intense stimulus, consistent with metacontrast masking (e.g., Alpern 1952) in that saccadic velocity diminishes toward the end of the saccade and the now-slowly-moving eye receives greater magnitudes of illumination distributed over smaller areas generating greater suppression. So for the high-intensity stimulus, by 100-msec duration, with about 65 msec of light in one retinal location following saccadic termination, the postsaccadic stimulus is sufficient to eliminate the smear from perception, and what is seen is a narrow line again, exactly as predicted from the mechanism based on metacontrast with stationary stimuli.

Thus, each and every stimulus in a visual field, at every location and at every intensity, carries its own inhibitor of presaccadic and during-the-saccade visible persistence, suppressing the smear before it can be perceived (backward masking), pushing the presaccadic view into a memory store that allows the postsaccadic local sign shift to be operated on by the extraretinal signal via cancellation. Nearly all of SSV (saccadic suppression of visibility) is a result of such visual inhibition from spatiotemporal stimulation to the retina (see E. Matin 1974, 1976; Riggs et al. 1982; and particularly Greenhouse & Cohn 1991, for a treatment of a smaller portion of SSV due to a lessening of sensitivity resulting from a mix of extraretinal factors); about 80% of SSD (saccadic suppression of displacement; Bridgeman et al. 1975) is a result of inhibition by extraretinal factors related to the eye movement (see Li & Matin 1997, for summary). So, saccadic suppression, both SSV and SSD, is central to a useful set of functions that allows perception to be stable in the presence of saccades. The neural basis for suppression of the presaccadic and during-the-saccade stimulation in the spatiotemporal distribution of stimulus energy is similar to that which generates sharpening of the perception of auditory frequency in the cochlea and of spatial localization of stimulation from the Békésy cochlear model applied to the skin (von Békésy 1967, 1968) as well as possessing a significant role in contour perception (Ratliff et al. 1958; Ratliff 1965).

6.2.6 Influences of visual context on saccadic localization

With a temporal paradigm similar to the one in Figs. 6.1 and 6.2, employing horizontal 8-deg-long saccades in the absence of a background, Honda (1993, Fig. 3) measured localization errors similar to those in Fig. 6.2. The pattern of errors did not depend on the horizontal location of the standard target in the psychophysical measurements. However, he also discovered that when a steady background was present the pattern depended significantly on the horizontal location of the standard target (Honda, Fig. 2). Increasing distance of the standard target from the fixation point produced a reduction of the positive error (see Fig. 6.2), virtually eliminating it, and increased the size of the negative error in response to stimulation during the saccade. In effect, this was manifested as a significant reduction in the differences in distance between apparent horizontal locations (measured by matching) of the standard targets during the saccade relative to the differences both prior to and following the saccade. Employing 20-deg-long saccades, Ross et al. (1997) confirmed Honda's result with the background, labeling the main result as "compression." Awater and Lappe (2004) subsequently confirmed Honda's result that compression required visual references and did not occur in their absence.

Thus, visual context is an important modulator of spatial localization. As shown in the first experiment (by E. Matin 1968) to examine the influence of visual context on spatial localization in the presence of saccades, even a continuously present single point of light exerts a major influence on localization (E. Matin et al. 1969; Matin & Matin 1972; Matin 1976b; Fig. 6.5). The legend in that 1972 article provided a caveat worth repeating. The results presented there "indicate the necessity of eliminating confounding contextual influences in experiments in which the attempt is to measure the extraretinal influences alone" (Matin & Matin 1972, p. 363). However, Ross et al. (1997) and Morrone et al. (1997) have presented a model with which they are able to obtain good fits to their compression data that make use of an extraretinal component and a compression function as simultaneously operating, additive, time-varying contributors to the apparent visual direction of flashes in the temporal neighborhood of saccadic eye movements in the presence of stationary visual context.

6.2.7 A hybrid EEPI and the effect of eye position

Whether an inflow or outflow signal provides the source for EEPI has remained controversial at least since a plausible physiological basis for inflow – the stretch receptor – was discovered (Sherrington 1898, 1918). The original outflow theory suggesting that EEPI resulted from a corollary discharge accompanying the command to turn the eyes had been proposed much earlier by von Helmholtz (1866/1963; for relevant reviews that also discuss some matters beyond our most immediate concerns see Matin 1972, 1976a, 1986; Steinbach 1987; Andersen et al. 1997; Colby & Goldberg 1999; Donaldson 2000; Schlag & Schlag-Rey 2002). Figure 6.6 displays the results of an experiment with saccades (Li 1989; Li & Matin 1992a) that requires contributions by both inflow and outflow components and

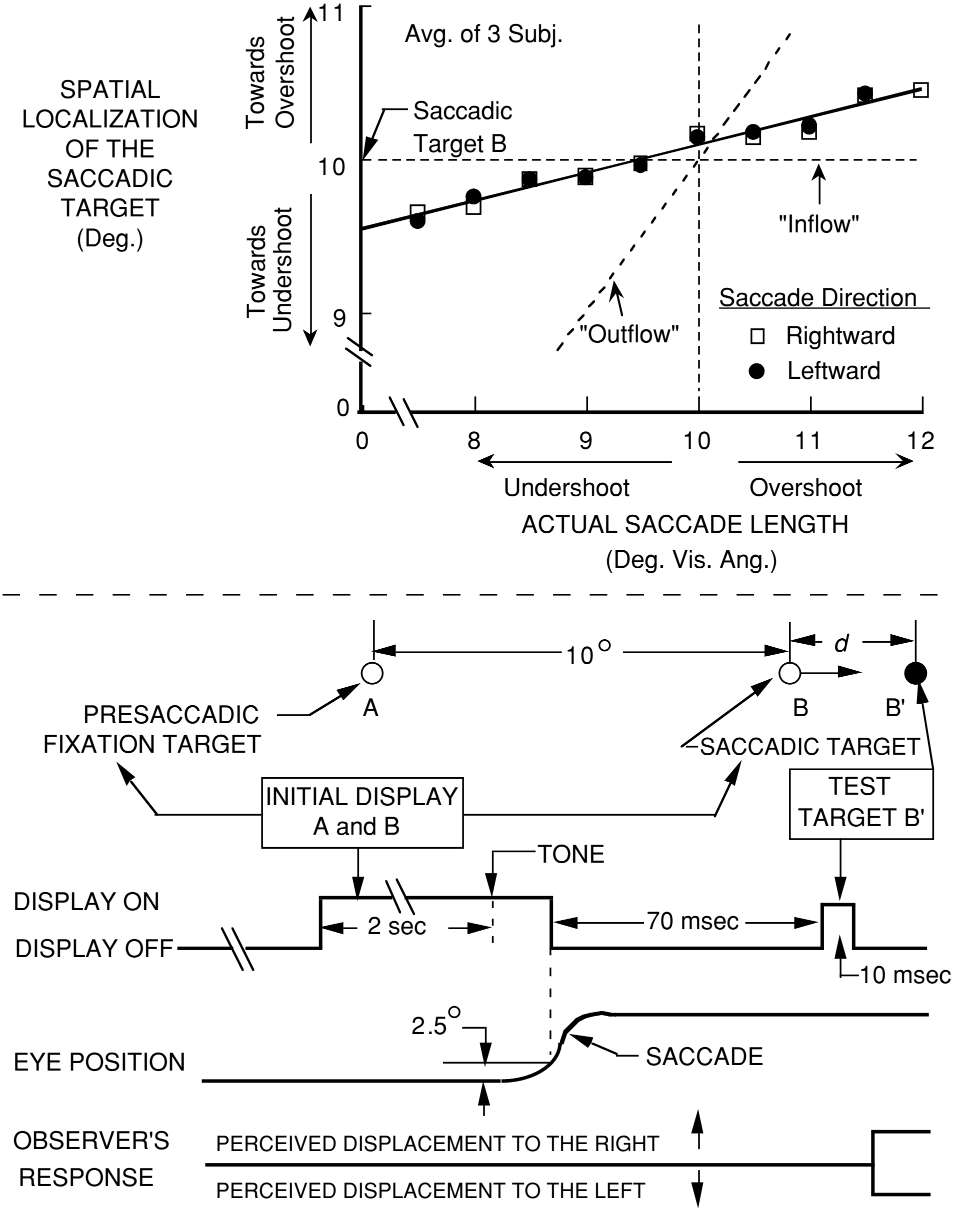


Fig. 6.6 The subject viewed the initial display and, following a tone, performed a saccade from the presaccadic fixation target (A) to the saccadic target (B). When the eye reached 2.5 deg into the saccade, a trigger from the eye-movement monitor turned the initial display off; 70 msec later a variably located test target (B') was flashed for 10 msec. The data for each S were separated into ten equally separated bins over the 4.5-deg range of saccade lengths, and a separate PSE was calculated for each bin for each of the three subjects; the average PSEs are displayed as a function of the deviation of the eye from the 10-deg saccadic target. Also displayed are the slope of 1.00 line (dashed diagonal) as the outflow locus, and the slope of 0.00 line (at the ordinate value of 10 deg) as the inflow locus. The average slope of the straight lines drawn through the subject's data is 0.193 indicating a contribution of 81% by inflow and of 19% by outflow (Li 1989; Li & Matin 1992a).

cannot reasonably be explained with only one. Here subjects reported the location of a test flash relative to a saccadic target in an experiment where parameters were chosen, as in the earlier experiments (Fig. 6.1), so as to avoid critical constraints from visual persistence. On each trial, 2 sec following initiation of the initial display, a tone instructed S to execute a 10-deg saccade; 2.5 deg into the saccade, a trigger from the eye movement monitor turned off the initial display and, 70 msec later, generated a 10-msec, variably located test flash. The saccades varied in length from about 2.5 deg shorter than the 10-deg attempt to 2.0 deg longer, and the accuracy of perceptual localization for the test flash was systematically related to the eye position at the termination of the saccade in a way that clearly required the useful operation of EEPI. The relation between the perceptual and oculomotor errors was linear, with a shallow slope that averaged 0.19 for the three Ss (Fig. 6.6). These results differ markedly from the 1965 experiment (Fig. 6.1(a)) where a retinotopic result was obtained for the saccadic target. Because visible persistence decreases markedly with increased flash duration (Fig. 6.4), the saccadic target was no longer present as a persisting visual image at the time of test flash presentation in Fig. 6.6 but only in longer-term memory (see Section 6.2.3), thus permitting the useful contribution by EEPI. Thus, the difference is consistent with the persistence explanation and earlier results. Here EEPI can be treated quantitatively as a consequence of a weighted combination of two components in a hybrid mechanism in which inflow contributes 81% and outflow the remaining 19% (note the inflow and outflow loci in Fig. 6.6). The fact that localization changes only slightly with the position of the eye in the orbit (less than 0.5-deg error in the PSE for a 2.5-deg error in eye position) demonstrates this dominance of inflow.

6.2.8 The cancellation mechanism

The assumption of a cancellation mechanism has been basic to all previous treatments of egocentric space perception throughout the more than 150-year literature (at least since von Helmholtz 1866/1963; also see Sperry 1950; von Holst & Mittelstaedt 1950; von Holst 1954). We've employed it as a basis for analysis above (Matin 1972, 1982, 1986). This model may be simply represented as

$$\mathrm{PSE} = \mathrm{RI} - \mathrm{EEPI}, \tag{1}$$

where the PSE represents the physical location that the observer matches to a visual target, RI is retinal location based on retinal local signs in which the spatial metric is given to perception by the structure of the visual nervous system, and EEPI is superimposed on that basis independently of the parametric details of stimulation. Thus, turning the eye horizontally by X degrees produces an X-degree retinal image shift (RI), and in an environment in which all objects are stationary, the EEPI signal ($= -X$) is assumed to be equal and opposite ($\mathrm{EEPI} = -\mathrm{RI}$), leading to cancellation in a subsequent comparator stage that does the algebra and leaves perceptual localization (i.e., the PSE) of all objects in the visual field unchanged. Its intuitive simplicity is based on the assumption of linear additivity between the two input processes (RI and EEPI) and derives from the geometric fact that the eye movement–driven shift of distance at the retina is essentially linear with change in eye position.

The model in Equation (1) is a simplification that does not show the involvement of the visual factors we have emphasized above: visual persistence, metacontrast, and visual latency. With the three variables in Equation (1) treated as functions of time, the representation in Equation (1) was shown to fail in the earliest experiments with saccades involving flashes (Matin & Pearce 1965; Matin et al. 1969, 1970) and was in need of further detailed development. Nevertheless, its intuitive appeal has resulted in its continuing to provide an important basis for most subsequent psychophysical work to the present time.

Recently, Pola (2004, 2007) has presented a mathematical development that treats experimental results for visual localization of flashes in the presence of saccadic eye movements emphasizing the contributions of persistence and visual latency. His model weights a time-varying theoretical extraretinal signal by a retinal signal whose temporal properties were chosen to be consistent with known durations of visual persistence and neural delays. The parameter choices were in part based on some well-known aspects of the temporal frequency characteristics of the visual nervous system for sinusoidally modulated flicker, and he incorporated this front-end modeling in a linear systems framework that also makes use of some quantitative properties of a back-end plant for generating saccadic eye movements. In his formulation, the integration of the moment-by-moment product of the extraretinal and retinal signals yields a "psychophysical extraretinal signal" that was employed in neural simulations for several variations of his model. Pola (2004) found good fits to the simulated results for single flashes before, during, and after saccadic eye movements (data as in Sections 6.2.2 and 6.2.3) for cases in which this psychophysical extraretinal signal is sluggish relative to the speed of the eye during the saccade as well as when the speed of the psychophysical extraretinal signal matches the time course of the change in eye position. The full model is shown to be a further development of the cancellation model in which eye position is subtracted from the psychophysical extraretinal signal to yield perceived target location. That it is a sophisticated version of the simple model in Equation (1) is made clear by noting that the eye position in the subtraction and the retinal signal are essentially identical when signs are taken into account appropriately. Pola (2007) fitted the model to the two-flash experiments with a variable interflash interval (Sogo & Osaka [Section 6.2.4]). For both the single-flash and two-flash results he obtains a surprising conclusion: the psychophysical extraretinal signal is confined to the postsaccadic period.

Pola's model is an interesting and novel approach. It brings the use of the extraretinal signal in providing perceptual stability to perisaccadic visual localization into the intellectual home that has led to significant advances in our contemporary understanding of temporal processing in vision and visual perception during the past century. Here it provided a basis for treatments of the relations between temporal resolution, temporal integration, and intensity discrimination across different adaptation levels, and for the de Lange function (e.g., Graham & Kemp 1938; de Lange 1954, 1958; Matin 1962, 1968; Brown 1965; Hood & Finkelstein 1986). It will remain for future work to further examine this approach to saccadic localization.

6.3 Beyond the saccade: egocentric spatial localization and sensorimotor behavior in the steady state

EEPI's function in the time period around a voluntary saccade is an important piece of the story regarding spatial constancy. However, a much broader picture is obtained by examining the large and significant influences of spatial induction normally generated by the orientation of the visual field on egocentric space perception and visually guided sensorimotor behavior. Demonstrating this influence is best achieved by changing the orientation of the entire visual field, a manipulation not readily available in most laboratory environments, and so the effects of such changes have not been frequently examined. The classic work by Witkin and Asch (Asch & Witkin 1948a,b; Witkin & Asch 1948a,b; Witkin 1949) with a roll-tilted room and the rod-and-frame situation opened the door on the study of these aspects of egocentric spatial orientation (see also Wertheimer 1912; Koffka 1935). It demonstrated the dominance of spatial induction by a roll-tilted visual field over gravitational factors mediated by the body-referenced mechanism[1] on the perception of verticality in the face of strong belief to the contrary (Gibson & Mowrer 1938). The work below describes some of the further developments and generalizations involving (a) the perception of elevation, (b) connections between the perceptions of verticality and elevation, (c) the basis for induction in oriented straight lines, and (d) influences of spatial induction on visually guided sensorimotor behavior demonstrating common influences of illusion-producing stimulation with egocentric space perception.

6.3.1 Modification of EEPI by paralysis

The major influences of spatial induction on the perception of elevation were first discovered in psychophysical experiments with five human subjects whose extraocular muscles were partially paralyzed under curare, substantially reducing transmission capability at the neuromuscular junction in a normal laboratory environment (Matin et al. 1982, 1983). Three facts are of greatest interest in the present context.[2]

First, the world looked essentially normal to the paretic subjects; things appeared at the correct places and orientations. However, as soon as all room illumination was extinguished, the subject saw a small (20′ diameter), dim, red fixation target at true eye level slowly descend to a position near the (invisible) floor. Normal illumination immediately restored the target's appearance to eye level again. This light/dark response sequence could be repeated as often as desired.

---

[1] The term "body-referenced mechanism" was introduced (Matin & Fox 1989) to refer to the combination of all extraretinal influences on the perception of interest, including extraretinal eye position information, extraretinal head orientation information (including information regarding the head relative to the body and the head relative to gravity), other effects of gravity on the body, pressure cues from the surfaces of the body, joint receptors, and the vestibular organ; it includes, in addition, the basic local sign information from the visual target employed to measure the discrimination itself.

[2] When attempting to execute a voluntary saccade to a specified goal, the paretic eye fell short of its goal, and a second saccade or series of saccades ensued that reached the goal with the following limitation: the total extent of possible gaze variation was considerably reduced by an amount approximately equal to the magnitude of visual mislocalization in the given dimension; the reduction was dose-dependent; initial saccade length was scaled to the reduction in total gaze extent (Matin et al. 1983).

Second, in otherwise complete darkness, large, steady-state mislocalizations of perceived elevation (VPEL: visually perceived eye level) and perceived laterality (VPSA: visually perceived straight ahead) occurred, measured by settings of a small target. These reached ±20 deg from true eye level and ±20 deg from the median plane, respectively, and were systematically related to the magnitude of the reduced-range, steady-state, voluntary deviations of the eye in the orbit relative to an orbital zero (Fig. 6.7), but were unrelated to head pitch.

Third, because the auditory system lies behind the blood–brain barrier, it was not influenced by curare, and auditory/visual matches in the horizontal dimension were mislocalized identically in illumination (Fig. 6.7(c,d)) and darkness, demonstrating the continued operation of a significantly modified extraretinal signal in normal illumination as well as in darkness.

The dramatic changes in egocentric localization with illumination demonstrated the greater potency of the visual field relative to influence from the body-referenced mechanism (although this difference is modified at high-g: Chelette et al. 1995; Li et al. 2001). The changes in darkness caused by change in transmission at the neuromuscular junction strongly supported an outflow basis for EEPI.[3] By passively pressing on the eye of normal subjects, Stark and Bridgeman (1983) created visual mislocalizations consistent with those for the outflow interpretation for the paralyzed eye. Along with a subsequent reinterpretation of the consequences of eyepress they report a small contribution of inflow (Bridgeman & Stark 1991). The evidence for major outflow and minor inflow effects in the paralyzed eye and eyepress cases appears to be in conflict with the evidence for major inflow and minor outflow in the saccade case (Section 6.2.7). However, the conflict is eliminated, although not satisfactorily explained, by noting that the paralyzed eye effects are measured with an essentially steady eye, whereas the effects with saccades involve transients.

6.3.2 Visual induction, the dominant influence in egocentric space perception

The experiments with paralyzed observers had raised a significant question: Why was the visual perception of elevation in the curarized state, although so badly mislocalized in darkness, normal in illumination? There are two main possibilities: (1) Consistency of the orientation of the illuminated visual field with the direction of gravity led to normalization of visual perception in the illuminated field. (2) The influence of visual input under illumination was combined with the abnormal EEPI or overrode it, generating the perception of elevation independently of any consistency between the orientation of the visual input and the direction of gravity.

Confirmation of the second possibility was obtained on normal observers with the aid of a pitchroom (for description see legend, Fig. 6.8(a)) (Matin & Fox 1986, 1989, 1990; Stoper & Cohen 1989). Consistency between visual and gravitational inputs was only a special case. Systematic variation of the pitch of the observer's visual environment generated

---

[3] As suggested by the hybrid model (Matin 1976, 1986), the possibility of modification by curare of signals from gamma efferents leading to spindles in extraocular muscle complicates the inference of purely outflow control from the paralyzed eye.

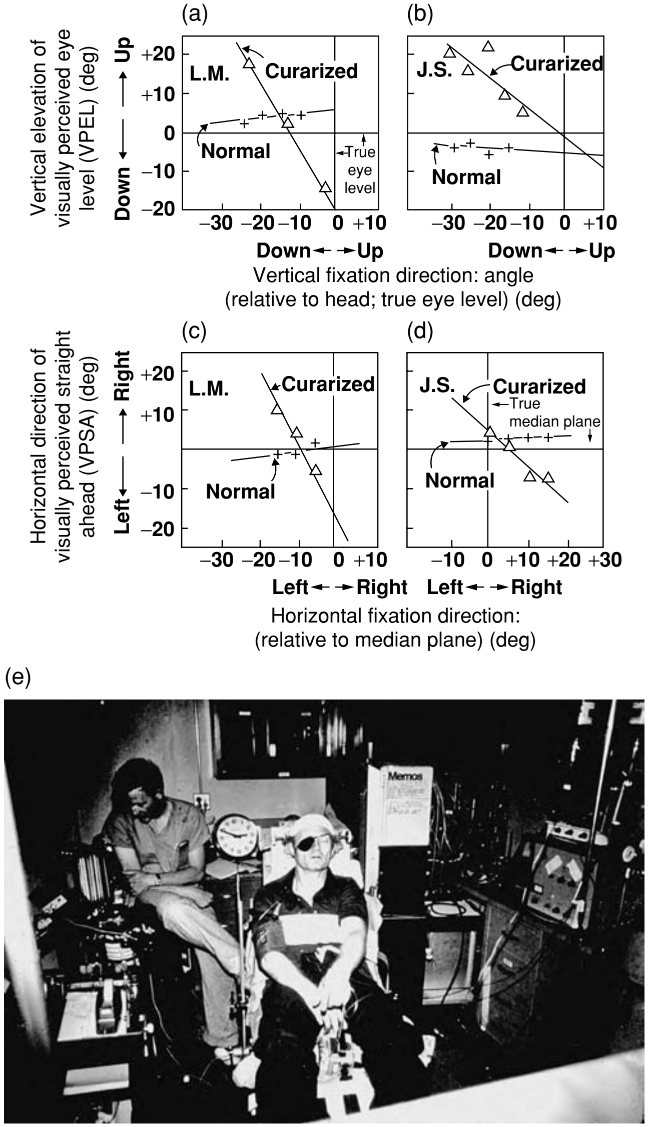


Fig. 6.7 Each of five subjects was continuously administered d-tubocurarine systemically to reduce the efficiency of the neuromuscular junctions of the extraocular muscles. This produced large errors in spatial localization of visual targets in otherwise complete darkness: (a,b). Each of two subjects set the elevation of a target to appear at eye level (VPEL setting) either in the normal (noncurarized) state or in the curarized state. The ordinate displays the error in visual angle units as a deviation from true eye level as a function of the elevation of the eye in the orbit (abscissa). (c,d) Each of the same two Ss set the horizontal location of a target to appear in the median plane ("straight ahead") in either the normal or curarized state. The ordinate displays the error in visual angle units as a deviation from the true median plane as a function of the horizontal position of the eye from the true median plane (abscissa). (e) Subject LM attempting to point to a visual target; due to the curare-induced systemic weakness, it was necessary for the right hand to assist the left (pointing) hand. The anesthesiologist (Edwards) is in the background (adapted from Matin et al. 1982).

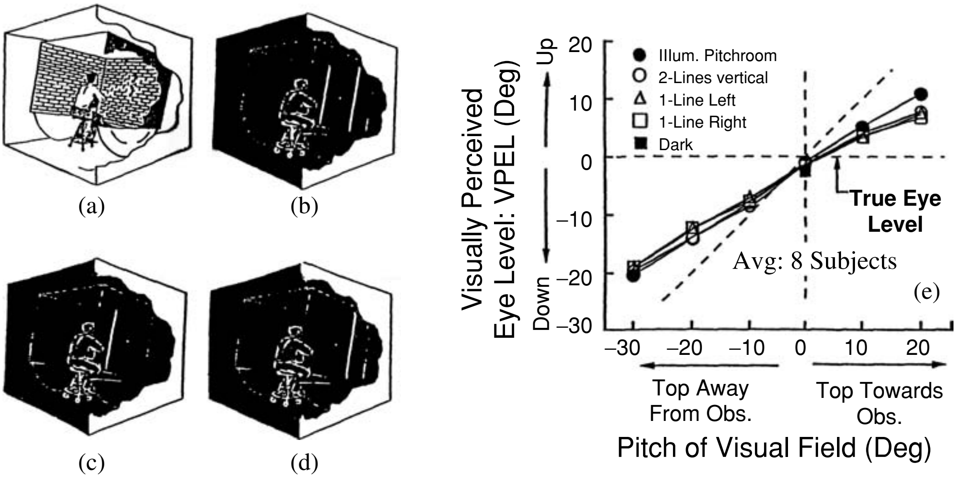


Fig. 6.8 Cutaway sketches of the pitchroom (pitched topbackward) inside the erect exterior room of the building. The rectangular structure is mounted on the base of two semicircular rockers resting on the exterior floor. A black cloth (not shown) was draped over the observer's shoulders and hung between the chinrest and the shelf of the pitchroom extending along the latter's horizontal extent; it prevented the observer from seeing any part of her body or the exterior room's floor. The walls and ceiling of the pitchroom along with the black cloth constituted the observer's entire visual field. The pitchroom's pitch was adjustable between 45 deg topbackward and 30 deg topforward (see Matin & Fox 1989 for further detail). The subject reported on the elevation of a small target within the median plane relative to eye level projected on the wall facing the subject; the settings of VPEL were either made in the variably pitched pitchroom (a), or in the completely dark pitchroom with only (b) the two long variably pitched inducing lines visible, or (c) only the left or (d) the right variably pitched inducing line visible. The data (e) display VPEL as a function of the pitch of the visual field (data from Matin & Li 1995).

large near-linearly related changes in VPEL. With the illuminated pitchroom set 40 deg topbackward (Fig. 6.8(a)) VPEL averaged 25.4 deg below true eye level (below the subject's hip); topforward pitch produced VPEL settings above true eye level, reaching 15.5 deg above true eye level with 25-deg pitch. If VPEL was determined by pitchroom orientation exclusively, the slope of the near-linear VPEL-versus-pitch function should have been 1.00. However, its 0.63 slope indicated that the pitchroom orientation (V) shared control with the extraretinal influence of the body-referenced mechanism (B) in a weighted average:

$$\mathrm{VPEL} = \mathrm{k_V V} + \mathrm{k_B B} \tag{2}$$

with $\mathrm{k_V} + \mathrm{k_B} = 1, \mathrm{k_V} = .63$, and $\mathrm{k_B} = .37$.

6.3.3 Identical retinal orientations induce changes in perceptions of elevation (VPEL) and verticality (VPV)

When we asked what there was about a pitched visual field that led to the dramatic effects on the perception of elevation we entered a geometric world filled with surprises.

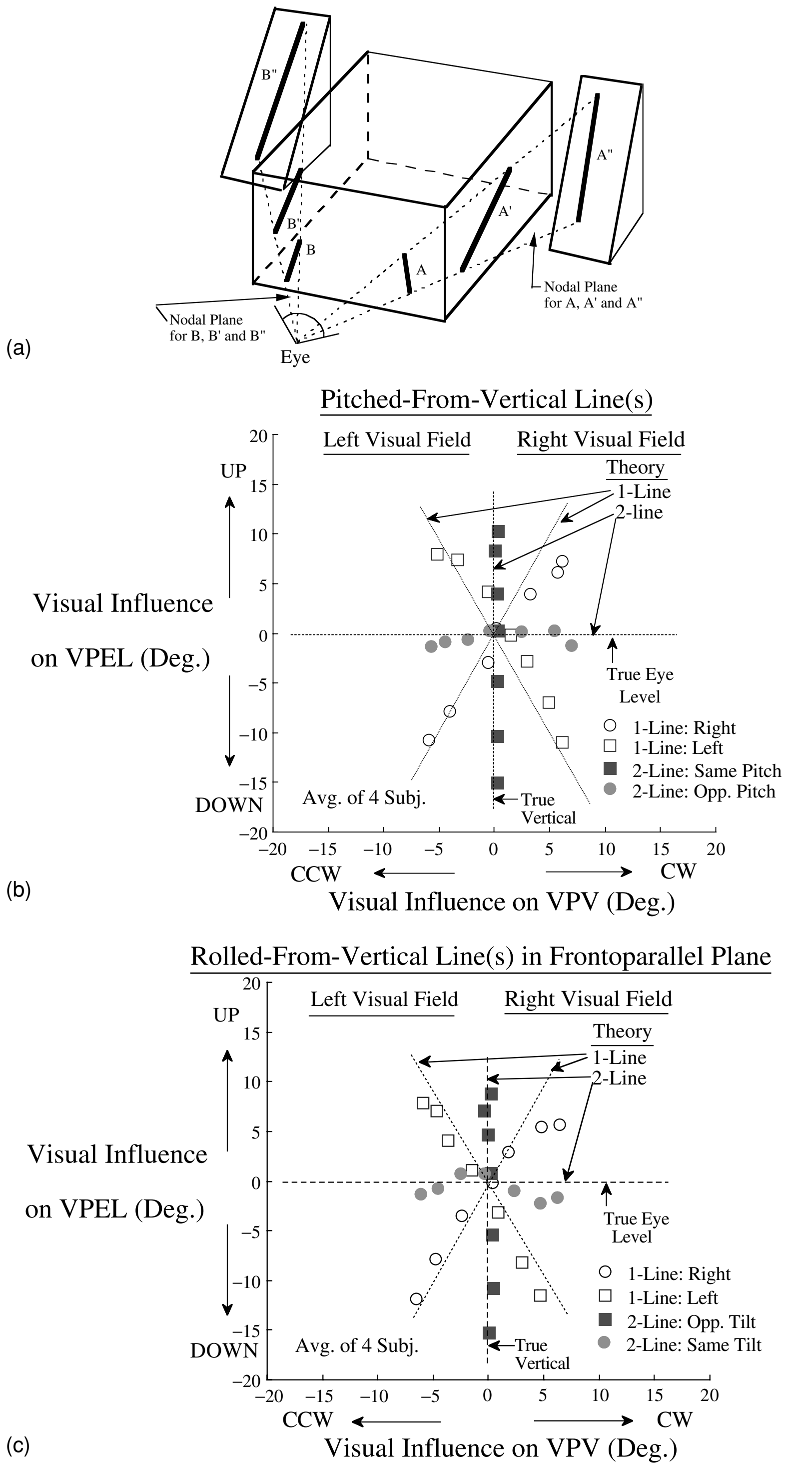


Fig. 6.9 (a) Differently oriented lines from three differently oriented planes in space lying within the same nodal plane produce stimulation of identical retinal orientations: (A, A′, and A″; B, B′, and B″);

First surprise: A single, long pitched-from-vertical line in total darkness (Fig. 6.8(c,d)) generated an effect on VPEL only 15% less than the entire well-illuminated, fully structured visual field in Fig. 6.8(a), and the effect of two such lines together (Fig. 6.8(b)) was only 10% less (Fig. 6.8(e)); horizontal lines produced only very small effects (Matin & Li 1994a).

Second surprise: (a) Because visual pitch is generated by rotating a frontoparallel plane or line around a horizontal axis in the erect observer's frontal plane, whereas visual roll is generated by rotating the same plane or line around an orthogonal horizontal axis in the observer's median plane, and (b) because visual pitch and visual roll produce different effects on egocentric space perception (as noted above, change in the elevation perceived as eye level [VPEL] by visual pitch; change in the frontal-plane orientation perceived as vertical [VPV] by visual roll), a very different fundamental basis in line orientation might be anticipated for VPEL than for VPV. However, this naive expectation fails. Individual lines from pitched and from rolled planes that lie in the same nodal triangle (Fig. 6.9(a)), A″ and A (or B″ and B), strike the same retinal orientation and exert the same influence on VPEL whether originating from a pitched or rolled plane (Matin & Li 1994b; Li & Matin 1996), emphasizing that the key to the influence of an individual line on VPEL is its ***retinal orientation***, not the physical plane from which it may arise.

But two other surprising results do hold here: **Third surprise:** Not only do all lines in a given nodal triangle (Fig. 6.9(a)) exert the same effect on VPEL, but they all also exert the same effect on VPV (Matin & Li 1994b, 1995). Thus, stimulation of a given retinal orientation influences both the perceptions of elevation ***and*** verticality (Figs. 6.9(b), 6.9(c)). **Fourth surprise:** (a) Whereas two parallel rolled lines stimulating bilaterally–symmetrically centered retinal orientations exert the same influence on VPV, they exert equal and opposite influences on VPEL. (b) Whereas two bilaterally symmetric lines stimulating bilaterally–symmetrically-centered retinal orientations exert the same influence on VPEL, they exert equal and opposite influences on VPV. The identity of effects of pitched-from-vertical lines and roll-tilted lines are displayed in the identity of the results in the two graphs of Fig. 6.9 (Matin & Li 1994b,c, 1995, 1996). In each graph a point plots the average VPEL on the ordinate and the average VPV on the abscissa for the identical

---

Fig. 6.9 (*cont.*)
the two different nodal planes shown stimulate bilaterally symmetric retinal orientations. Note that parallel lines A″ and B″ are pitched from vertical but that A and B, bilaterally symmetric for the viewing eye, are rolled from vertical. (b) Average results of psychophysical measurements of VPEL (visually perceived eye level) and VPV (visually perceived vertical) for which the VPEL (ordinate) and VPV (abscissa) values shown as coordinates of a given point were measured with the same pitched-from-vertical inducer; VPEL and VPV were measured in different sessions. The inducers were either a single line in the right or left visual fields ("1-Line Right", unfilled circles; "1-Line Left", unfilled squares), or a pair of lines where the two members of the pair possessed either the same pitch or equal and opposite pitches ("2-Line Same Pitch", filled squares; "2-Line Opp Pitch", filled circles). (c) Same as (b) employing rolled-from-vertical lines from an erect frontoparallel plane; each rolled-from-vertical line in (c) stimulated the same retinal orientation as one of the pitched-from-vertical lines in (b); the same four subjects were employed in (b) and (c). (Matin & Li 1996).

inducing line(s) at one of seven different retinal orientations. Thus, both (a) and (b) are directly displayed in Fig. 6.9(b) and 6.9(c). The equivalence of the influences of the left and right visual fields on elevation (VPEL) is manifested in the left–right symmetry in both Figs. 6.9(b) and 6.9(c), and the equivalence of influences on verticality (VPV) in the up–down symmetry in both Figs. 6.9(b) and 6.9(c). Thus, line orientation mediated by retinal orientation not only provides the basis for the VPEL results in the full pitchroom but also provides the basis for the VPV results in the rod-and-frame situation (Li & Matin 2005b,c).

With a great deal of success, a neuromathematical model predicts the VPEL response to variation of orientation and length of 1-line inducers, the response to 2-line inducers of all combinations of orientation and lengths, and to the fully structured visual field of the pitchroom (Matin & Li 2001). Generalizations of these influences have also been discovered that have shown that the global orientation of a series of points (Li & Matin 1992b; Matin et al. 2006) and an array of short parallel lines orthogonal to their individual local orientations induce effects on VPEL of the same order as do the individual lines, suggesting connections with the perception of form and texture (Shavit et al. 2004, 2005).

6.3.4 Influence of eye elevation and generalized cancellation

For a given pitchroom orientation (±20 deg), VPEL is invariant over a ±20 deg vertical range of eye position (Li & Matin 1993; Matin & Li 1995), demonstrating that the combination of RI and EEPI maintains their relation (cancellation) when a pitch-induced error of fixed magnitude is added. This leads to a generalization of Equation (1): With VPEL as the PSE,

$$\mathrm{VPEL} = \mathrm{RI(ep)} - \mathrm{EEPI(ep)} + \mathrm{k_B B} + \mathrm{k_V V}. \tag{3}$$

Equation (3) thus represents the influences of the internal (EEPI) and external (V) influences in a single framework along with visual stimulation from the target of the discrimination (RI). V equals visual pitch or equivalent pitch of the pitched or rolled lines (e.g., Fig. 6.9(a)), and RI and EEPI are functions of eye position (ep). Equation 3 applies to the steady state in the pitchroom as well as to the transient (perisaccadic) situation and the curarized state (Matin et al. 1982, 1983; Matin & Fox 1986, 1989; Matin & Li 1992, 1995, 2001; Li et al. 2001; Cohen et al. 1995). This makes it clear that the pitchroom illusion arises as a consequence of having broken into the fundamental mechanism that stabilizes the perception of elevation against changes in eye position.

6.3.5 Spatial induction and hand-to-body distance dependence of visually guided sensorimotor manual behavior

The **fifth and sixth surprises** are centered on visually guided manual behavior. In spite of large errors in perceived elevation and perceived verticality that are systematic with pitch

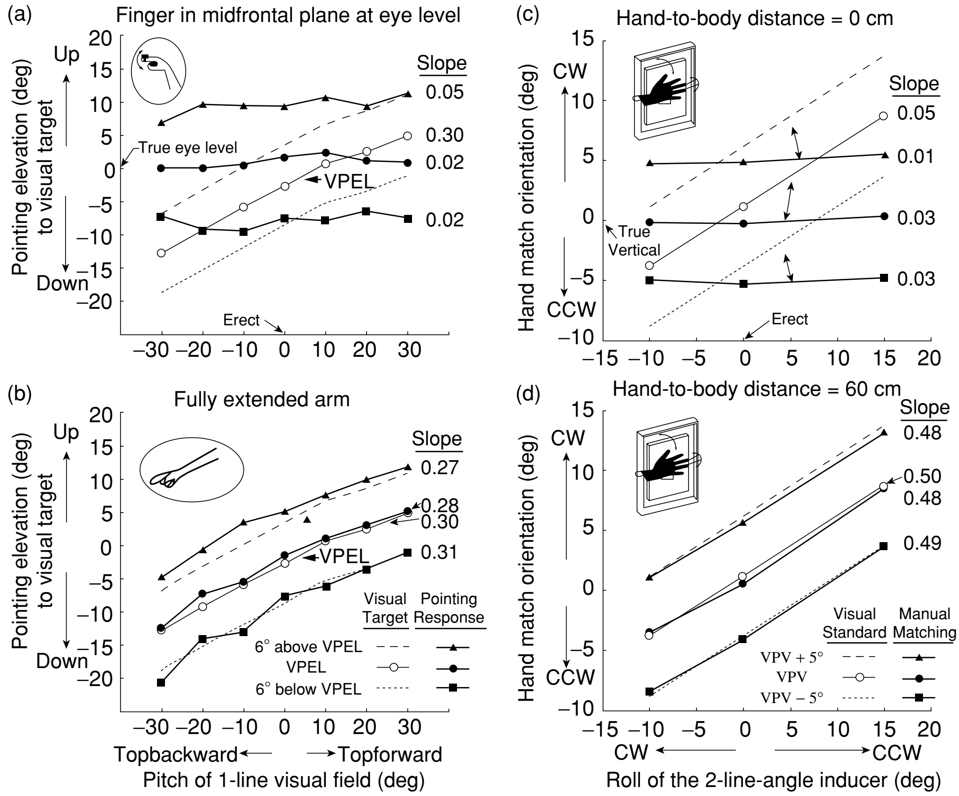


Fig. 6.10 Each of four subjects set a small visual target to appear at eye level (VPEL; open circles in (a) and (b)) under induction by a single dim, 50-deg-long × 4.5′-wide line (25 deg horizontal eccentricity) set at each of seven pitches in otherwise total darkness. In (a) with the hand in the midfrontal plane and fixed at true eye level, the subject pointed (filled circles) the index finger at the visual target whose elevation was at the elevation previously set to VPEL. In (b) the manual settings were made with the finger at the end of the fully extended arm (filled circles). In separate sessions the subject pointed to the visual target that was 6 deg above (filled triangles) or 6 deg below (filled squares) VPEL in the presence of the inducer; the dashed and dotted lines are 6 deg above and below the VPEL locus (open circles) (Li & Matin 2005a). (c) and (d) Each of six subjects first set the orientation of a short line in the median plane to appear vertical (VPV; open circles) under induction by a two-line right-angle inducer (“half-a-square”) at each of three physical orientations in the frontoparallel plane (10 deg ccw, erect, or 15 deg cw) in otherwise total darkness. In (c) and (d) manual settings to match the perceptual VPV settings were made with the hand in the midfrontal plane and at 60 cm from the midfrontal plane, respectively (filled circles). The subject average of the rolled-from-vertical manual settings from the midfrontal plane did not depart significantly from veridicality regardless of the orientation of the visual target that changed systematically by about 13 deg, but with the hand fully extended (60 cm) from the midfrontal plane, full accuracy was attained in matching the target that appeared to be vertical (VPV), although VPV varied by about 13 deg. Entirely consistent results were attained in (c) and (d) with the target of the match set 5 deg cw (filled triangles) or 5 deg ccw (filled squares) from VPV; the dashed and dotted lines are 5 deg cw and ccw from the VPV locus (Matin et al. 2004; Li et al. 2008).

and roll (Figs. 6.10(a–d); Li & Matin 2005a; Li et al. 2008), with a fully extended arm, subjects were able to point accurately at the elevation of a visual target under induction by a single, long, variably pitched line (Fig. 6.10(c)) and to set the orientation of the hand to accurately match the orientation of a tilted line under induction by a tilted 2-line right-angle frame (Fig. 6.10(d)), both in otherwise total darkness. However, with the hand close to the body (in the midfrontal plane; 0 cm hand-to-body distance), manual errors changed linearly with pitch or roll of the inducer. Thus, the elevation of S's finger point was invariant when pointing to the visual target that appeared at eye level (at VPEL), although the target's physical elevation was changed systematically with visual pitch (Fig. 6.10(a)), and the orientation of S's hand was invariant when set to match the orientation of the target line that appeared erect (at VPV), although the physical orientation of the target line changed systematically with the inducing frame's roll tilt (Fig. 6.10(b)). Further experiments (Li & Matin 2005a; Li et al. 2008) have shown that the influences at the two extremes of hand-to-body distance are parts of variations in manual accuracy that are linearly graded with hand-to-body distance.

These large hand-to-body distance-dependent modulations of manual accuracy to perceptually mislocalized visual targets make it clear that the dissociation between "perception" and "action" systems reported for an illusion of size contrast (Aglioti et al. 1995; Milner & Goodale 1995; Haffenden & Goodale 2000) is not general (see Li & Matin 2005a, footnotes 2 and 3, for summaries and references to the large, extended controversy). Further, the attempt to align the perception/action dichotomy with the separation of ventral and dorsal streams (Milner & Goodale 1995), although plausible, is not sufficient to account for the results in Fig. 6.10.

The results on manual behavior in Fig. 6.10 disclose a connection between perception and action that is different with the hand in near space and far space. This manifests useful manual accuracy to distant objects where accuracy is needed in dealing with the object regardless of perceptual inaccuracy, and it manifests a complete lack of visual influence with the hand close to the body where visual influence that would, for example, put the hand 20 deg below the face would seriously hamper normal manual behavior, such as putting food in one's mouth or scratching one's ear.

Acknowledgments

The research described in this article was supported by NSF Grant BCS-06–16654, by grant EY 10534 from NEI, NIH, and by earlier grants from the Spatial Orientation Program of AFOSR.

References

Aglioti, S., DeSouza, J. F. X., & Goodale, M. A. (1995). Size-contrast illusions deceive the eye but not the hand. *Curr Biol* **5**: 679–685.

Alpern, M. (1952). Metacontrast: historical introduction. *Am J Optometry* **29**: 631–646.

Alpern, M. (1953). Metacontrast. *J Opt Soc Am* **43**: 648–657.

Andersen, R. A., Snyder, L. H., Bradley, D. C., & Xing, J. (1997). Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Ann Rev Neurosci* **20**: 303–330.

Asch, S. E., & Witkin, H. A. (1948a). Studies in space orientation. I. Perception of the upright with displaced visual fields. *J Exp Psychol* **38**: 325–337.

Asch, S. E., & Witkin, H. A. (1948b). Studies in space orientation. II. Perception of the upright with displaced visual fields and with body tilted. *J Exp Psychol* **38**: 455–477.

Awater, H., & Lappe, M. (2004). Perception of visual space at the time of pro- and anti-saccades. *J Neurophysiol* **91**: 2457–2464.

Bockisch, C. J., & Miller, J. M. (1999). Different motor systems use similar damped extraretinal eye position information. *Vision Res* **39**: 1025–1038.

Boucher, L., Groh, J. M., & Hughes, H. C. (2001). Afferent delays and the mislocalization of perisaccadic stimuli. *Vision Res* **41**: 2631–2644.

Bowen, R. (1981). Latencies for chromatic and achromatic visual mechanisms. *Vision Res* **21**: 1483–1499.

Bowen, R., Pola, J., & Matin, L. (1974). Visual persistence: effects of flash luminance, duration, and energy. *Vision Res* **14**: 295–303.

Breitmeyer, B. G., & Ganz, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression and information processing. *Psychol Rev* **83**: 1–36.

Bridgeman, B., Hendry, D., & Stark, L. (1975). Failure to detect displacement of the visual world during saccadic eye movements. *Vision Res* **15**: 719–722.

Bridgeman, B., & Stark, L. (1991). Ocular proprioception and efference copy in registering visual direction. *Vision Res* **31**: 1903–1913.

Brown, J. L. (1965). Flicker and intermittant stimulation. In C. H. Graham (ed.), *Vision and Visual Perception* (251–320). New York: Wiley & Sons.

Cai, R. H., Pouget, A., Schlag-Rey, M., & Schlag, J. (1997). Perceived geometrical relationships affected by eye-movement signals. *Nature* **386**: 601–604.

Campbell, F. W., & Wurtz, R. H. (1978). Saccadic omission: why we do not see a grey-out during a saccadic eye movement. *Vision Res* **18**: 1297–1303.

Chelette, T., Li, W., Esken, R., & Matin, L. (1995). Visual perception of eye level (VPEL) under high g while viewing a pitched visual field. *Annual Meeting of Aerospace Medical Association*: **A55**.

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.

Colby, C. L., & Goldberg, M. E. (1999). Space and attention in parietal cortex. *Ann Rev Neurosci* **22**: 319–349.

Dassonville, P., Schlag, J., & Schlag-Rey, M. (1992). Oculomotor localization relies on a damped representation of saccadic eye displacement in human and nonhuman primates. *Vis Neurosci* **9**: 261–269.

Dassonville, P., Schlag, J., & Schlag-Rey, M. (1993). Direction constancy in the oculomotor system. *Curr Dir Psychol Sci* **2**: 143–147.

Dassonville, P., Schlag, J., & Schlag-Rey, M. (1995). The use of egocentric and exocentric location cues in saccadic programming. *Vision Res* **35**: 2191–2199.

de Lange, H. (1954). Relationship between critical flicker frequency and a set of low-frequency characteristics of the eye. *J Opt Soc Amer* **44**: 380–389.

de Lange, H. (1958). Research into the dynamic nature of the human fovea-cortex systems with intermittant and modulated light. I. Attenuation characteristics with white and coloured light. *J Opt Soc Amer* **48**: 777–784.

Donaldson, I. M. L. (2000). The functions of the proprioceptors of the eye muscles. *Philos Trans R Soc Lond B* **355**: 1685–1754.

Efron, R. (1970). Effect of stimulus duration on perceptual onset and offset latencies. *Perception & Psychophysics* **8**: 231–234.

Gibson, J. J., & Mowrer, O. H. (1938). Determinants of the perceived vertical and horizontal. *Psychol Rev* **45**: 300–323.

Graham, C. H., & Kemp, E. H. (1938). Brightness discrimination as a function of the duration of the increment in intensity. *J Gen Physiol* **21**: 635–650.

Greenhouse, D. S., & Cohn, T. E. (1991). Saccadic suppression and stimulus uncertainty. *J Opt Soc Amer* **8**: 587–595.

Haffenden, A. M., & Goodale, M. A. (2000). Independent effects of pictorial displays on perception and action. *Vision Res* **40**: 1597–1607.

Hallett, P. E., & Lightstone, A. D. (1976a). Saccadic eye movements towards stimuli triggered by prior saccades. *Vision Res* **16**: 99–106.

Hallett, P. E., & Lightstone, A. D. (1976b). Saccadic eye movements to flashed targets. *Vision Res* **16**: 107–114.

Hansen, R. M., & Skavenski, A. A. (1985). Accuracy of spatial localization near the time of saccadic eye movements. *Vision Res* **25**: 1077–1082.

Honda, H. (1989). Perceptual localization of visual stimuli flashed during saccades. *Perception & Psychophysics* **45**: 162–174.

Honda, H. (1990). Eye movements to a visual stimulus flashed before, during, or after a saccade. In M. Jeannerod (ed.), *Attention and Performance*, vol. 13. Hillsdale, NJ: Erlbaum.

Honda, H. (1991). The time courses of visual mislocalization and of extraretinal eye position signals at the time of vertical saccades. *Vision Res* **31**: 1915–1921.

Honda, H. (1993). Saccade-contingent displacement of the apparent position of visual stimuli flashed on a dimly illuminated structured background. *Vision Res* **33**: 709–716.

Honda, H. (1999). Modification of saccade-contingent visual mislocalization by the presence of a visual frame of reference. *Vision Res* **39**: 51–57.

Hood, D. C., & Finkelstein, M. A. (1986). Sensitivity to light. In K. Boff, L. Kaufman, & J. Thomas (eds.), *Handbook of Perception and Human Performance* (vol. I, ch. 5, 5-1–5-66). New York: Wiley.

Kahneman, D. (1968). Methods, findings, and theory in studies of visual backward masking. *Psychol Bull* **70**: 404–425.

Koffka, K. (1935). *Principles of Gestalt Psychology*. New York: Harcourt Brace.

Kolers, P. A. (1962). Intensity and contour effects in visual masking. *Vision Res* **2**: 277–294.

Kolers, P. A., & Rosner, B. (1960). On visual masking (metacontrast): dichoptic observation. *Am J Psychol* **73**: 2–21.

Li, W. (1989). A quantitative study of the accuracy and the precision of the visual perception of direction immediately after voluntary saccades. Ph.D. dissertation, Columbia University.

Li, W., Dallal, N., & Matin, L. (2001). Influences of visual pitch and yaw on visually perceived eye level (VPEL) and straight ahead (VPSA) for erect and rolled-to-horizontal observers. *Vision Res* **41**: 2873–2894.

Li, W., & Matin, L. (1992a). Visual direction is corrected by a hybrid extraretinal eye position signal. *Ann New York Acad Sci* **656**: 865–867.

Li, W., & Matin, L. (1992b). Rotation in depth of linear arrays of points systematically influences egocentric localization. *Bull of the Psychonomic Soc* **30**: 453.

Li, W., & Matin, L. (1993). Eye and head position, visual pitch, and perceived eye level. *Inv Ophth & Vis Sci* **34**: 1311.

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. (1997). Saccadic suppression of displacement: separate influences of saccade size and of target retinal eccentricity. *Vision Res* **37**: 1779–1797.

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 Res* **45**: 533–550.

Li, W., & Matin, L. (2005b). Visual induction of visually perceived vertical (VPV) by 1-line and 2-line roll-tilted and pitched visual fields. *Vision Res* **45**: 2037–2057.

Li, W., & Matin, L. (2005c). The rod-and-frame effect: the whole is less than the sum of its parts. *Perception* **34**: 699–716.

Li, W., Matin, E., Bertz, J., & Matin, L. (2008). A tilted frame deceives the eye and the hand. *J Vision* **8**(16): 18, 1–16.

Lotze, H. (1886). *Outline of Psychology*. Translated and edited by Ladd, G. T. Boston: Ginn.

Matin, E. (1974). Saccadic suppression: a review and an analysis. *Psychol Bull* **81**: 899–917.

Matin, E. (1976). Saccadic suppression and the stable world. In R. A. Monty & J. W. Senders (eds.), *Eye Movements and Psychological Processes* (113–119). New York: Erlbaum.

Matin, E. (1982). Saccadic suppression and the dual mechanism theory of direction constancy. *Vision Res* **22**: 335–336.

Matin, E., Clymer, A., & Matin, L. (1972). Metacontrast and saccadic suppression. *Science* **178**: 179–182.

Matin, E., Matin, L., & Pola, J. (1969). The intermittant light illusion and constancy of visual direction during voluntary saccadic eye movements. *Psychonomic Society* (Abstract).

Matin, L. (1962). Fourier treatment of some experiments in visual flicker. *Science* **136**: 983–985.

Matin, L. (1964). Measurement of eye movements by contact lens techniques: analysis of measuring systems and some new methodology for three-dimensional recording. *J Opt Soc Am* **54**: 1008–1018.

Matin, L. (1968). Critical duration, the differential luminance threshold, critical flicker frequency, and visual adaptation: a theoretical treatment. *J Opt Soc Am* **58**: 404–415.

Matin, L. (1972). Eye movements and perceived visual direction. In D. Jameson & L. Hurvich (eds.), *Handbook of Sensory Physiology* (331–380). Heidelberg: Springer.

Matin, L. (1976a). A possible hybrid mechanism for modification of visual direction associated with eye movements – the paralyzed-eye experiment reconsidered. *Perception* **5**: 233–239.

Matin, L. (1976b). Saccades and extraretinal signal for visual direction. In R. A. Monty & J. W. Senders (eds.), *Eye Movements and Psychological Processes* (203–204). New York: Erlbaum.

Matin, L. (1982). Visual localization and eye movements. In A. Wertheim, W. A. Wagenaar, & H. Leibowitz (eds.), *Tutorials on Motion Perception* (101–156). New York: Plenum.

Matin, L. (1986). Visual localization & eye movements. In K. Boff, L. Kaufman, & J. Thomas (eds.), *Handbook of Perception and Human Performance* (vol. I, ch. 20, 20-1–20-45). New York: Wiley.

Matin, L., & Bowen, R. W. (1976). Measuring the duration of perception. *Perception & Psychophysics* **20**: 66–76.

Matin, L., & Fox, C. R. (1986). Perceived eye level: elevation jointly determined by visual pitch, EEPI, and gravity. *Invest Ophthalmol Vis Sci (Suppl.)* **27**: 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 Res* **29**: 315–324.

Matin, L., & Fox, C. R. (1990). Erratum: Visually perceived eye level and perceived elevation of objects: linearly additive influences from visual field pitch and from gravity. *Vision Res* **30**: I.

Matin, L., & Kibler, G. (1966). Acuity of visual perception of direction in the dark for various positions of the eye in the orbit. *Percept Mot Skills* **22**: 407–420.

Matin, L., & Li, W. (1992). Visually perceived eye level: changes induced by a pitched-from-vertical 2-line visual field. *J Exp Psychol Hum Percept Perform* **18**: 257–289.

Matin, L., & Li, W. (1994a). The influence of the orientation of a stationary single line in darkness on the visual perception of eye level. *Vision Res* **34**: 311–330.

Matin, L., & Li, W. (1994b). Spatial summation among parallel lines across wide separations (50°): spatial localization and the great circle model. *Vision Res* **34**: 2577–2598.

Matin, L., & Li, W. (1994c). Mirror symmetry and parallelism: two opposite rules for the identity transform in space perception and their unified treatment by the Great Circle Model. *Spat Vis* **8**: 469–489.

Matin, L., & Li, W. (1995). Multimodal basis for egocentric spatial localization and orientation. *J Vestib Res* **5**: 499–518.

Matin, L., & Li, W. (1996). Similarities between 3D color and egocentric orientation spaces. *Inv Ophth & Vis Sci* **37**/3: S519.

Matin, L., & Li, W. (2001). Neural model for processing the influence of visual orientation on visually perceived eye level (VPEL). *Vision Res* **41**: 2845–2872.

Matin, L., Li, W., & Bertz, J. W. (2004). Distance-contingent accuracy of manual matches to line orientations misperceived under the 2-line rod-and-frame illusion. *J Vis* **4**(8): 380a.

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) [Abstract]. *J Vis* **6**: 972.

Matin, L., & Matin, E. (1972). Visual perception of direction and voluntary saccadic eye movements. In J. Dichgans & H. Bizzi (eds.), *Cerebral Control of Eye Movements and Motion Perception* (358–368). Basel, Switzerland: Karger.

Matin, L., Matin, E., & Pearce, D. G. (1969). Visual perception of direction when voluntary saccades occur. I. Relation of visual direction of a fixation target extinguished before a saccade to a flash presented during the saccade. *Perception & Psychophysics* **5**: 65–80.

Matin, L., Matin, E., & Pola, J. (1970). Visual perception of direction when voluntary saccades occur: II. Relation of visual direction of a fixation target extinguished before a saccade to a subsequent test flash presented before the saccade. *Perception & Psychophysics* **8**: 9–14.

Matin, L., Matin, E., Pola, J., & Bowen, R. (1971). Relative visual direction of two flashes presented at different times or intensities during a voluntary saccade – retinal constraints in the operation of extraretinal signals. Presented at EPA.

Matin, L., & Pearce, D. G. (1964). Three-dimensional recording of rotational eye movements by a new contact-lens technique. In W. E. Murry & P. F. Salisbury (eds.), *Biomedical Sciences Instrumentation* (79–95). New York: Plenum Press.

Matin, L., & Pearce, D. G. (1965). Visual perception of direction for stimuli flashed during voluntary saccadic eye movements. *Science* **148**: 1485–1488.

Matin, L., Pearce, D. G., Matin, E., & Kibler, G. (1966). Visual perception of direction in the dark: roles of local sign, eye movements and ocular proprioception. *Vision Res* **6**: 453–469.

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* (243–262). New York: Springer.

Miller, J. M. (1996). Egocentric localization of a perisaccadic flash by manual pointing. *Vision Res* **36**: 837–851.

Miller, J. M., & Bockisch, C. J. (1997). Where are the things we see? *Nature* **386**: 550–551.

Milner, A. D., & Goodale, M. A. (1995). *The Visual Brain in Action*. New York: Oxford University Press.

Morrone, M. C., Ross, J., & Burr, D. C. (1997). Apparent position of visual targets during real and simulated saccadic eye movements. *J Neurosci* **17**(20): 7941–7953.

Pola, J. (1973). The relation of the perception of visual direction to eye position during and following a voluntary saccade. Ph.D. dissertation, Columbia University.

Pola, J. (1976). Voluntary saccades, eye position, and perceived visual direction. In R. A. Monty & J. W. Senders (eds.), *Eye Movements and Psychological Processes* (245–254). New York: Erlbaum.

Pola, J. (2004). Models of the mechanism underlying perceived location of a perisaccadic flash. *Vision Res* **44**: 2799–2813.

Pola, J. (2007). A model of the mechanism for the perceived location of a single flash and two successive flashes presented around the time of the saccade. *Vision Res* **47**: 2798–2813.

Raab, D. H. (1963). Backward masking. *Psychol Bull* **60**: 118–129.

Ratliff, F. (1965). MACK BANDS: *Quantitative studies on neural networks in the retina*. San Francisco, CA: Holden-Day, Inc.

Ratliff, F., Miller, W. H., & Hartline, H. K. (1958). Neural interaction in the eye and the integration of receptor activity. *Ann NY Acad Sci* **74**: 210–222.

Riggs, L. A., Volkmann, F. C., Moore, R. K., & Ellicott, A. G. (1982). Perception of suprathreshold stimuli during saccadic eye movement. *Vision Res* **22**: 423–428.

Ross, L., Morrone, M. C., & Burr, D. C. (1997). Compression of visual space before saccades. *Nature* **386**: 598–601.

Schlag, J., & Schlag-Rey, M. (2002). Through the eye, slowly: delays and localization errors in the visual system. *Nat Rev Neurosci* **3**: 191–215.

Shavit, A. Y., Li, W., & Matin, L. (2004). Influences of global and local orientations of line segments on perceived eye level. 45th Annual Meeting of the Psychonomic Society, **9**: 112.

Shavit, A. Y., Li, W., & Matin, L. (2005). Spatial induction of changes in perceived elevation and verticality by global and local orientations of sets of lines. 5th Annual Meeting of Vision Sciences Society, 275.

Shebilske, W. L. (1977). Visuomotor coordination in visual direction and position constancies. In W. Epstein (ed.), *Stability and Constancy in Visual Perception*. New York: Wiley.

Sherrington, C. S. (1898). Decerebrate rigidity, and reflex coordination of movements. *J Physiol (Lond.)* **22**: 319–332.

Sherrington, C. S. (1918). Observation on the sensual role of the proprioceptive nerve supply of the extrinsic ocular muscles. *Brain* **41**: 332–343.

Sogo, H., & Osaka, N. (2001). Perception of relation of stimuli locations successively flashed before saccade. *Vision Res* **41**: 935–942.

Sogo, H., & Osaka, N. (2002). Effects of inter-stimulus interval on perceived locations of successively flashed perisaccadic stimuli. *Vision Res* **42**: 899–908.

Sperry, R. (1950). Neural basis of the spontaneous optokinetic response produced by vision inversion. *J Comp Physiol Psychol* **43**: 482–489.

Stark, L., & Bridgeman, B. (1983). Role of corollary discharge in space constancy. *Perception & Psychophysics* **34**: 371–380.

Steinbach, M. J. (1987). Proprioceptive knowledge of eye position. *Vision Res* **27**: 1737–1744.

Stoper, A., & Cohen, M. M. (1989). Effect of structured visual environments on apparent eye level. *Perception & Psychophysics* **46**: 469–475.

von Békésy, G. (1967). *Sensory Inhibition*. Princeton, NJ: Princeton University Press.

von Békésy, G. (1968). Mach- and Hering-type lateral inhibition in vision. *Vision Res* **8**: 1457–1466.

von Helmholtz, H. (1866). *Handbuch der Physiologischen Optik*, Leipzig: Voss. (English translation from 3rd edition 1925). In J. P. C. Southall (ed.), *A Treatise on Physiological Optics* (vol. 3). New York: Dover.

von Holst, E. (1954). Relation between the central nervous system and the peripheral organs. *British Journal of Animal Behavior* **2**: 89–94.

von Holst, E., & Mittelstaedt, H. (1950). Das reafferenz-prinzip: Wechselwirkungen zwischen zentralnervensystem und peripherie (The principle of reafference: interactions between the central nervous system and the peripheral organs). *Naturwissenscaften* **37**: 464–477.

Weisstein, N. (1972). Metacontrast. In D. Jameson & L. Hurvich. (eds.), *Handbook of Sensory Physiology* (vol. 7, no. 4, Visual Psychophysics). Berlin: Springer-Verlag.

Wertheimer, M. (1912). Experimentelle studien ü ber das sehen von bewegung, *Zeitschrift fur Psychologie und Physiologie der Sinnesorgane* **61**: 161–265.

Witkin, H. A. (1949). Perception of body position and of the position of the visual field. *Psychol Monograph* **63**(7): 1–46.

Witkin, H. A., & Asch, S. E. (1948a). Studies in space perception. III. Perception of the upright in the absence of a visual field. *J Exp Psychol* **38**: 603–614.

Witkin, H. A., & Asch, S. E. (1948b). Studies in space perception. IV. Further experiments on perception of the upright with displaced visual fields. *J Exp Psychol* **38**: 762–782.

7

Space constancy: the rise and fall of perceptual compensation

BRUCE BRIDGEMAN

Summary

Information about eye position comes from efference copy, a record of the innervation to the extraocular muscles that move the eye and proprioceptive signals from sensors in the extraocular muscles. Together they define extraretinal signals and indicate the position of the eye. By pressing on the eyelid of a viewing eye, the extraocular muscles can be activated to maintain a steady gaze position without rotation of the eye. This procedure decouples efference copy from gaze position, making it possible to measure the gain of the efference copy signal. The gain is 0.61; the gain of oculomotor proprioception, measured by a similar eye press technique, is 0.26. The two signals together sum to only 0.87, leading to the conclusion that humans underestimate the deviations of their own eyes and that extraretinal signals cannot be the mechanisms underlying space constancy (the perception that the world remains stable despite eye movements). The underregistration of eye deviation accounts quantitatively for a previously unexplained illusion of visual direction. Extraretinal signals are used in static conditions, especially for controlling motor behavior. The role of extraretinal signals during a saccade, if any, is not to compensate the previous retinal position but to destroy it. Then perception can begin with a clean slate during the next fixation interval.

7.1 Introduction

All visual information arrives in the brain through the retinas, whose images are displaced with each eye movement. Yet we perceive a stable visual world. How does the brain accomplish the stabilization? This problem defines one of the fundamental accomplishments of visual perception: space constancy, the perception that the world remains fixed even as the eyes scan across it. Perceiving a stable visual world establishes the platform on which all other visual function rests, making possible judgments about the positions and motions of the self and of other objects. The perception seems paradoxical.

A half-century ago, it seemed that the problem of space constancy had been solved; the basic mechanism was known, and it remained only to find the physiological substrate and clean up the details. The solution was a signal emanating from motor areas of the brain to inform the visual system about when and where the eyes had moved. At the time of an

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

eye movement this signal could be subtracted from the shift of the retinal image, achieving visual space constancy. Because the idea requires that the visual centers receive a copy of the neural efference to the eye muscles, it is now called efference copy. The efference copy is an extraretinal signal (Matin 1972) affecting vision but not originating from the retina. It is also called "outflow" because a signal flows out from the oculomotor centers to compensate for retinal image motion (Teuber 1960).

The solution was a long time coming; ideas about something coming out of the brain, complementing what was coming in, go back to the ancient Greeks (Grüsser 1986a). Their idea was very different, though; for Aristotle, some sort of energy emanated from the eyes to allow vision to take place. Thinking that the eyes of animals seeming to glow in the dark was a visible manifestation of such emanations, they concluded that the emanations interacted with objects in the world to mediate vision. Arab scholars followed this interpretation. Needless to say, such ideas did little to advance vision science.

7.2 Early formulations

The efference copy idea originated with the physiologist Charles Bell (1823/1974), who had already discovered the Bell/Magendie law of separated afferents and efferents in the spinal cord. At about the same time Purkinje (1825) also described the idea, apparently independently. Both descriptions are based in part on perceptions that occur when the side of an eye is pressed with a finger. If the eye is pressed in darkness with an afterimage on the retina, no motion of the afterimage is perceived. An active eye movement, though, will result in apparent movement of the afterimage. Experience with a real image is just the reverse – it appears to move when the eye is pressed but does not move with a voluntary eye movement. These four observations could be explained if an active eye movement elicited an extraretinal signal to compensate for eye movement, but the eye press did not.

The failure of afterimage movement with the eye press in darkness would be inevitable, for the afterimage would remain fixed on the retina while the eye press did not elicit an extraretinal signal. The movement of the afterimage with an eye movement in otherwise dark surroundings could be explained only by an efference copy, for only the efference copy changes in this condition. A normal eye movement in a normal environment would not elicit apparent motion because the retinal image motion would be matched by the efference copy. But the eye press in a normal environment would elicit apparent motion because the resulting retinal image motion would not be compensated by the extraretinal signal.

The four conditions are neatly explained with a single theory summarized in Table 7.1. The conditions in bold type result in space constancy, either with both efference copy and active eye movement or with neither efference copy nor active eye movement. The other two conditions represent failure of space constancy because of a mismatch between efference copy and image movement.

Both Bell and Purkinje went further to conclude that gaze movement signals canceled retinal image displacements to achieve space constancy. Somewhere in the brain, signals

Table 7.1

	Retinal Image Motion	No Retinal Image Motion
Efference copy	**Normal eye movement**	Afterimage with saccade
No efference copy	Eye press in normal field	**Eye press in darkness**

representing change in position of a retinal image were subtracted from signals representing change in oculomotor innervation.

7.2.1 The founding of physiological optics

For more than a century after this, efference copy was the major mechanism assumed to mediate space constancy. Hering (1861/1990) further asserted that one should obtain compensation for voluntary eye movements but not for involuntary movements such as vestibular afternystagmus. These are the involuntary eye movements induced by continuing vestibular activity following sustained head rotation, accompanied by feelings of dizziness and perceived motion of the visual world. The breakdown in space constancy occurs because the eye movements are driven directly by the vestibular system in a three-neuron arc that does not activate the normal outflow signal.

The dominance of outflow mechanisms in explaining space constancy was assured by von Helmholtz in his *Physiological Optics* (1866/1962), then and now the most influential work in the field. He expanded the empirical base for outflow theories with observations of neurological patients collected by Albrecht von Graefe. These patients had muscle pareses so that they could not use part of their oculomotor fields. When they attempted to look into the paretic field, the world seemed to jump in the direction of the intended movement, and pointing to a target in that direction went too far in the direction of the intended movement.

In analyzing these observations, von Helmholtz extended the efference copy idea to include sensorimotor coordination as well as perception. The patient has two facts to evaluate, for example, in pointing with a gaze that is paralyzed for movements to the right:

1. I am looking toward the right.
2. There is an image on my fovea (the fixational area of the retina).

The reasonable conclusion is that there is an image to the patient's right, although due to failure of the eyes to move, the gaze has actually remained straight ahead. Von Helmholtz called this reasoning an "unconscious inference," analogous to the processes of formal logic but executed effortlessly and without training. Pointing too far in the direction of the paretic field ("past pointing"), to the right in this case, shows that the patient has no information from eye muscle proprioception or any other source that might inform him of the actual gaze position. It is only the intended gaze position that affects perception and action. Von

Helmholtz called the intention to change gaze position a "Willensanstrengung," an effort of will.

The explanation is similar for the perception of a jump of the world in the direction of an intended eye movement. Before the intended jump there is an image on the fovea and a Willensanstrengung straight ahead. After the intended jump the eyes have not moved because of the paresis, but the Willensanstrengung is now directed toward the right, and the same image is still on the fovea. The conclusion is that the image has now jumped to the right, because eye position (as reported by Willensanstrengung) has changed, but the retinal image position has not.

Von Helmholtz also gave four observations in normal subjects supporting his outflow theory. First, moving the eye passively results in apparent motion; second, moving the eye passively does not result in apparent motion of an afterimage; third, image displacement is compensated in normal saccades (space constancy); fourth, adaptation to displacing prisms transfers intermanually. However, perhaps because von Helmholtz saw his eye movement signal as related to the will, he did not analyze it mathematically. Mach (1906), another physicist–physiologist, made that step by hypothesizing that a neuronal copy of oculomotor efference sums algebraically with the retinal signal to yield a position of viewed objects relative to the head. The first flow diagrams to define the concept came from von Uexküll (1920, 1928), who foreshadowed later mathematical analyses by differentiating efference copy from eye muscle proprioception and describing the consequences of each.

7.2.2 Mathematical theories and the modern era

Two papers appeared independently in 1950 that defined efference copy theory for the next generation. In fact, the phrase "efference copy" first appeared in an article in German by Erich von Holst and Horst Mittelstaedt (1950) as "Efferenzkopie." This was an empirical article, describing the results of inverting the head of the blowfly *Eristalis* by rotating its neck 180 deg and holding it there with a bit of wax (the blowfly has a very flexible neck). Von Holst and Mittelstaedt observed that the fly would circle continually. When the fly was in darkness, though, its locomotion seemed normal. With light restored, the fly would circle either in the original direction or in the opposite direction at random.

These results were explained with the assumption that the fly monitored the output of its locomotor system. The results compared that output with the retinal flow field (because the *Eristalis* eye is fixed to the head, the locomotor system is also the oculomotor system). The copy of locomotor efference, the Efferenzkopie, would be subtracted from the retinal signal to stabilize locomotion by negative feedback. Inverting the head converted the negative feedback to positive feedback – a random nudge in one direction would feed back a signal to "correct" in the same direction. That would result in a further deviation in the same direction, and continuous circling would result.

Von Holst and Mittelstaedt also contributed an engineering flow diagram and algebraic analysis, with the efference copy exactly canceling the afferent retinal signal (Fig. 7.1). This seminal article also introduced the terms "exafference," a retinal motion signal resulting

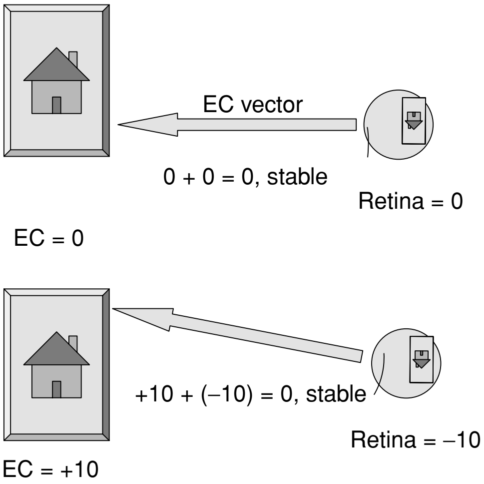


Fig. 7.1 Perceptual compensation by efference copy (EC). The EC, expressed as innervations to the six extraocular muscles controlling each eye, can be conceived as a vector along the line of sight. Top: The eye in primary position receives an image of the world. Bottom: The eye moves up 10 deg. As it does so, the EC indicates a 10-deg upward deviation and the retinal image shifts by 10 deg. The two signals of equal magnitude cancel, resulting in constant apparent position.

from motion of objects in the world, and "reafference," a retinal motion signal resulting from the organism's movements.

Sperry (1950) made similar observations in a fish whose eye he inverted surgically. He concluded that his fish's normal swimming in the dark excluded the possibility of brain or nerve damage and introduced the term "corollary discharge" to identify the efferent signal.

These articles formalized the quantitative compensation idea that had dominated physiology and psychology for more than a century. The new evidence offered for the idea was motor rather than sensory in nature, an emphasis that would prove important in the coming decades, although some speculations about perception were made.

7.3 Problems with compensation theories

Although compensation theories completely dominated thinking about space constancy to this point, there had always been problems with them. Considerations from control theory, which had made rapid progress during World War II, made these problems clear.

7.3.1 Spatial problems

First, the efference copy is a feedforward, a signal that informs the brain of where the eyes ought to be rather than where they actually are. As such it cannot be exact – it should drift with time, and it is not corrected when it is in error. Yet the perception of space constancy is perfect – the world does not appear to jump in the slightest when the eyes move. To the average person, the idea that the world should jump with each saccade seems bizarre at best. If perception is rock solid, but the efference copy is not, something else must be supplementing the feedforward signal, and that something else might be all that is necessary to do the job.

Recognizing that efference copy could not be perfect, E. Matin (1974) proposed that saccadic suppression could mask the inevitable errors. It was known then that displacements of the entire visual world would not be detected perceptually if they occurred during saccadic eye movements (Wallach & Lewis 1965; Mack 1970); if the imprecision of efference copy was less than the displacement thresholds during saccades, space constancy could be maintained despite small mismatches of efference copy and retinal displacement.

Matin's solution was the best idea available at the time, but it didn't last long. The first parametric description of saccadic suppression of displacement showed that, at the optimal timing of image displacement and saccade, the perceptual threshold was about one-third as large as the saccade itself (Bridgeman et al. 1975). Clearly, any visual orientation mechanism that tolerated an error of one part in three had no idea where the visual world was and could support neither perceptual space constancy nor a reasonable visual–motor calibration.

This result along with similar observations should have led to a capitulation of the efference copy theory, but it did not. The reason why is that a theory cannot be abandoned because of evidence; it can only be replaced by another theory, and none was at hand. There was a realization, though, that efference copy would not be the answer to the space constancy question.

7.3.2 Temporal problems

Before long, other problems with the efference copy theory began to surface. One of them began with the technique of reverse modeling, applying an output (behavior) to a linear model and running the equations backward to read the input (nerve signals to the muscles) necessary to drive the behavior. Applied to the oculomotor system, reverse modeling was able to clarify the motor signals that drive voluntary nystagmus, a rapid oscillation of the eyes that can be performed by a small proportion of otherwise normal people. The oscillations are small in amplitude, usually 3 deg or less, but high in frequency, up to 20 Hz in most cases. The resulting rotational accelerations of the eye are so great that the oculomotor driving signals can be generated only by the saccadic controller. Even though the movements have a nearly sinusoidal profile, they must be elicited by the pulse-step mechanism of saccades; the sinusoidal appearance is a result of temporal filtering by the oculomotor plant. All of this is relevant to the space constancy question because subjects experience oscillopsia, a back-and-forth fluttering of the visual world, during voluntary nystagmus. In short, space constancy breaks down.

But normally space constancy survives saccades, which are accompanied by saccadic suppression. What is going on? The possibility that small saccades do not elicit saccadic suppression was disproved by Nagle et al. (1980), who compared suppression during

Table 7.2

	Retinal Image Motion	No Retinal Image Motion
Space constancy	**Normal saccade**	**Nystagmus with afterimage**
No space constancy	Nystagmus in normal field	**Saccade with afterimage**

voluntary nystagmus to suppression during single voluntary saccades matching the amplitude of nystagmus in the corresponding subject. The suppression was virtually identical in both cases, demonstrating that saccadic suppression does not necessarily accompany space constancy. Further, an afterimage remained motionless during voluntary nystagmus (Table 7.2), showing that the changes of eye position failed to elicit changes in apparent position. Space constancy must use some other mechanism.

The voluntary nystagmus experiment showed that single isolated saccades were accompanied by space constancy, whereas rapidly alternating saccades of the same size were not. Perhaps the space constancy mechanism was still operating but could not keep up with rapidly alternating saccades of voluntary nystagmus. Grüsser et al. (1984) achieved a better temporal resolution of the constancy–frequency relationship in studies of the apparent movement of an afterimage with saccades in darkness. They asked subjects to make saccades from one loudspeaker to another, cued by tones from each speaker. After a bright light gave a lasting afterimage, saccades were performed in darkness. Grüsser et al. measured the subjects' estimates of the spatial separation of the afterimages when the eye was aimed at the left speaker versus the right speaker. As saccades became more frequent, the subjective separation of the afterimages became smaller and smaller, until at the highest saccade frequency (about 3.8 saccades/sec) the afterimage appeared to remain fixed in front of the subject. Space constancy had failed completely.

The result showed that voluntary nystagmus frequency was far higher than space constancy could handle and that perceptual compensation is quite slow. Even for intersaccade intervals well within the temporal range of saccades accompanying normal perception, the compensation was much smaller than the saccade amplitudes.

The temporal properties of space constancy were linked directly with efference copy a few years later in experiments exploiting the deceptively simple maneuver of pressing on the outer canthus of the eye. Explaining the method in these experiments requires a brief diversion into methodology.

The consequences of a gentle press on the outer edge of the eyelid have been misunderstood for centuries, since Purkinje's 1825 assumption that the press resulted in a passive eye movement, as reviewed previously. Von Helmholtz (1866) made the same assumption, that pressing on the eye moves it passively, and that the resulting apparent motion originates from retinal image movement without an efference copy. Two observations support this interpretation. First, the entire visual world appears to move in the direction opposite the eye press; and second, the eye of another person appears to move when it is observed during their eye press. The two observations are consistent with one another, but both are misinterpretations.

The apparent movement is based on the inference of motion of the retina, but the two kinds of motion are not necessarily linked. The real situation is easily demonstrated – simply pick a fixation target, then slowly press on the outer canthus of one eye while closing the other. You will find that you can hold your gaze on your fixation target, even while the entire visual world, fixation target and all, appears to move. This means that the retina is not moving at all with respect to the visual world.

If the motion is not coming from retinal slip, it must come from another source. That source originates with the successful effort of oculomotor tracking mechanisms, which cannot be turned off, in keeping the eye on the fixation target despite the eye press. The effort requires oculomotor innervation, and with it a change in efference copy. Far from demonstrating the effect of passive eye movement, the eye press demonstrates the effects of active compensation for oculomotor disturbance and shows that efference copy alone can drive the perception of motion.

The second misinterpreted observation is that the eye of another person performing an eye press appears to move. What the observer sees, however, is not an eye rotation but a lateral translation of the eye in the orbit. The eye is influenced by two rotational forces in opposite directions; one originates from the pressing finger producing a nasal-ward rotational force, and the other is an equal and opposite force generated by the lateral rectus muscle. The oculomotor innervation is driven by a retinal slip initiated from the finger but compensated by an involuntary optokinetic tracking system. Thus the two rotational forces cancel, and the eye does not rotate.

But each of these forces also introduces a translational component in the medial direction, the finger pushing the anterior part of the eye nasal-ward and the lateral rectus pulling the posterior part of the eye nasal-ward. The translational forces sum to move the eye several millimeters in the orbit, as measured experimentally (Stark & Bridgeman 1983). Because the cue that humans use to perceive movements of the eyes of others is the amount of sclera visible on the two sides of the iris, the translational motion is misinterpreted by observers as a rotation. The rotation of the occluded fellow eye, whose rotation is not canceled by the eye press, provides an objective measure of the forces applied.

Now the eye-press technique, which causes a deviation in efference copy without a change in retinal image position, can be applied to the problem of measuring the temporal aspects of efference copy. It was possible to use scleral search coils in both eyes simultaneously, and also to press on the eye without popping out the required scleral contact lens (the experiment is not for the fainthearted). Again the nonpressed eye is occluded so that its movements are measured in darkness. In this experiment, extending the static experiments of Stark and Bridgeman (1983), we pressed repeatedly on the viewing eye in a roughly sinusoidal pattern (Ilg et al. 1989). Adding to the complexity of the setup, a force transducer on the fingertip provided an objective record of the frequency and timing of the eye presses.

Replicating the work of Stark and Bridgeman (1983), we found that at low temporal frequencies the viewing eye does not rotate. Only the occluded eye rotates, under its occluder, revealing the compensatory oculomotor innervation; according to Hering's law (Hering 1868), that innervation affects both eyes equally. When we began pressing more rapidly on the eye, however, the compensation was no longer complete. At a rate of less than 1 Hz the occluded eye still rotated, but in addition the viewing eye rotated passively as it was repeatedly pressed and released. At the surprisingly low rate of 2 Hz, the occluded eye ceased its rotation completely, and only the viewing eye rotated in the passive manner that Purkinje and von Helmholtz would have predicted. Interpolation of our data implied that the oculomotor compensation system ceases to function at about 1–1.5 Hz. The implication is that any efference copy–based system that normally contributes to space constancy must cease to function at these relatively low rates, well within the bandpass of normal perceptual events.

By 1989, then, evidence from a number of directions was converging on the idea that efference copy could not be responsible for space constancy. Its action was too slow and its gain too low to support a perceptual compensation for eye movements. The theory continued to dominate, however, because no theory was available to replace it.

There was also a more qualitative sort of evidence that should have eliminated efference-based theories from consideration, but did not, again because of the lack of an alternative. One bit of evidence came from an experiment on saccadic suppression by Brune and Lücking (1969), who fed an eye movement signal into a mirror that moved an image with the eyes, but at variable gain (output/input). At low gains, when the image was moving one-tenth as far as the eye, the image appeared always to be stable, replicated by the findings of Bridgeman et al. (1975). But at a slightly higher gain, when the world as a whole continued to appear stable, "prominent objects" would seem to jump or jiggle with each saccade. The efference copy theories, however, do not allow the possibility that parts of the image can move relative to one another – the visual world is conceived as a monolithic object. The observation would seem to eliminate all efference copy and related theories in a single stroke.

There are technical reasons, however, why the Brune and Lücking experiment might have resulted in dissociations. The prominent objects might have been brighter than the background, for example, and therefore signals coding them would move through the visual system at a faster rate than signals from dimmer parts of the image. In a continuously moving environment, this might result in prominent objects being perceived in different locations than the context. Another possible artifact is that the prominent objects might have been fixated so that signals from them would course through the visual system in slower, high-acuity channels. Again, relative mislocalizations might result.

All of these possibilities were eliminated in a replication and extension of the study that used tesselations of a plane by the Dutch artist Maurits Escher as the stimulus materials (Bridgeman 1981). Escher used two repeated shapes that interlocked to completely cover a surface. For instance, devils and angels might tesselate a plane. Some subjects could selectively concentrate on just the angels, or just the devils, at will. All of those subjects saw slight movement of the attended figure while the "background" figure remained stable, at a near-threshold feedback gain from eye movement to image movement. Because this perception occurred with both figures, without any change in the stimulus, all image variables were controlled. Something was wrong with compensation theories – none of them could account for this result.

Another observation, one that had been known since the nineteenth century, is “autokinetic” motion of a small bright dot in darkness. Normally the visual world remains quite stable, but after a few seconds the small dot begins to appear to wander slowly through the dark field. No compensation theory can account for this observation, for these theories require an equal quality of constancy regardless of image properties. Space constancy in the efference copy theory requires an extraretinal signal that matches whatever comes in through the retina, regardless of structure or extent.

Autokinetic motion is now thought to originate from noise in the vestibular system affecting eye movements through a vestibulo-ocular reflex not registered in perception (Leibowitz et al. 1983). The noise drives the eye away from a target, and pursuit eye movements, which are registered in perception, are required to cancel the eye drifts. In a full field, however, the optokinetic reflex, which is also not registered in perception, can stabilize the field. The observations can be made consistent with efference copy theory only if one assumes that some kinds of eye movements are accompanied by an efference copy whereas others are not. It then becomes impossible for any brain mechanism comparing efference copy and retinal input to know what head-centered position to assign to the retinal input (Bridgeman 1995).

If the extraretinal signal theories have so many problems, what is the use of the extraretinal signals? An answer came from quantitative work on the gains of the efference copy and proprioception signals, building on a half-century of investigations of these signals. The work again exploited the static eye press technique, but with an additional twist. Pressing on the side of the viewing eye changed efference without changing gaze position, but pressing on the occluded eye should change only proprioception. The argument is that the occluded eye when pressed will be forced to rotate under the eyelid because the press does not result in any corrective signal from error feedback. If the proprioceptive signals from the two eyes are summed in the brain (as they must be according to Hering’s law), the resulting binocular gaze signal would equal half the deviation of the occluded eye (Fig. 7.2).

As infrared techniques had already been developed to monitor this eye position in darkness, the proprioception could be measured, and its effect on behavior could be assessed simultaneously by having subjects point to targets while eye press deviates the occluded eye. The situation in pressing the viewing eye is now more complicated because the perceptual changes will result from a combination of two signals working in opposite directions. Proprioceptive signals will come from the deviated, occluded eye, and altered efference copy will be driven by the active compensation for the press of the viewing eye. (This analysis was suggested by Wenshun Li.)

With these improvements in the eye press technique, it became possible to quantify gains of both outflow and inflow in normal observers in the same experiment. The internal signals could be recovered by algebraic rearrangement of the measured signals (Bridgeman & Stark 1991). Careful measurements of perceptual deviations with various magnitudes of eye press on the viewing or the occluded eye resulted in magnitudes of deviations that could

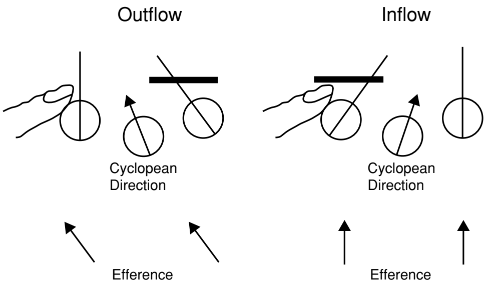


Fig. 7.2 Manipulation of outflow and inflow signals with monocular eye press in monocular viewing. Left: In the outflow condition the viewing eye is pressed, resulting in a compensatory efference (and efference copy (EC)). The cyclopean direction is half of the deviation of the occluded eye because the cyclopean signal is averaged with the undeviated viewing eye. Right: In the inflow condition the occluded eye is pressed, and it deviates mechanically because there is no visual error signal to correct gaze posture. Efference does not change. Again, cyclopean direction and proprioceptive signal equals half the eye's deviation.

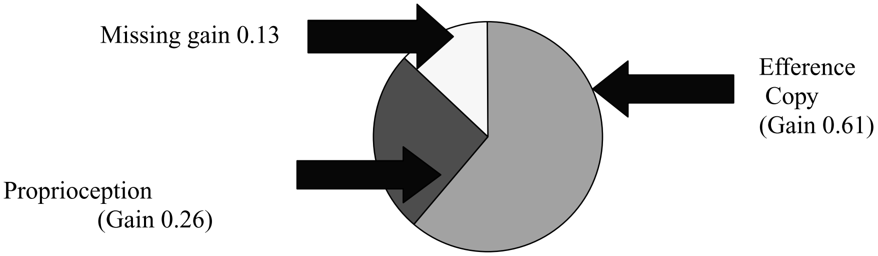


Fig. 7.3 Gains of efference copy and proprioception sum to 0.87, leaving an underregistration of 0.13 that is the origin of an eccentric-gaze illusion.

be used to recover the internal inflow and outflow signals. The resulting gains were 0.61 for outflow and 0.26 for inflow, a disturbing result because, even with perfect summation of the two gains, the brain would underestimate how far the eye really moved. There was a "missing" gain of 0.13 (Fig. 7.3). Where did it go?

The answer came from an earlier study of the perception of the deviation of an eccentric target from an observer's midline. Targets are perceived as more eccentric if their position is judged while looking at them through peripheral vision, with the eyes straight ahead, than if they are fixated with eccentric gaze first (Morgan 1978). This implies that eye deviations are registered as being smaller than the actual eccentricity of gaze. Quantifying these observations, Morgan drew a graph of her results but did not go further. It was a simple matter to calculate the gain from her graph, and the result was 0.13, precisely the "missing" gain from the eye press experiments. Outflow, inflow, and illusion gains sum to 1.00, closing the circle on the signals used in registering eye gaze position and the resulting perceptions.

As a result of this quantitative accounting for the eccentric-gaze illusion, we can draw several previously unrealized conclusions. First, inflow and outflow gains are summed in the brain's calculation of eye eccentricity. Two centuries of work on efference copy and proprioception led to the conclusion that efference copy dominates. We can now see that the reason for the apparent domination is that the efference copy gain is much higher than the proprioceptive gain, about 2.4 times greater. Thus, efference copy explains a wide range of results and clinical observations better than proprioceptive input does. Second, proprioceptive deviations are not compensated in eye posture. Presses on the occluded eye are passive, resulting in no oculomotor compensation.

The role of efference copy and of extraretinal signals generally, then, appears to be to inform the brain about static eye position during visual fixation, the time between saccades when the retina is transducing the visual world reliably. It does not support space constancy.

7.4 Modern alternatives for space constancy

Despite all of the difficulties with it, efference copy was discarded only recently as a mechanism for space constancy, with a new theory centered on a reanalysis of what information is carried over from one fixation to the next. The break came in 1992 when Kevin O'Regan asserted that it is not necessary to link successive images together – there need be no memory of the content of previous fixations because the information remains in the world and can be reacquired whenever the observer wants it. What the brain possesses is currently available retinal information and nothing more. This idea, that transsaccadic memory is in the world rather than in the brain, turned out to be too radical, but not by much.

Two years later another reanalysis appeared, along with a critique of previous theories (Bridgeman et al. 1994a,b). According to this analysis three information sources are traditionally used to achieve space constancy: proprioceptive inflow from eye muscles, efference copy outflow, and retinal information. The work reviewed previously, in addition to other physiological studies, converges on the conclusion that none of these sources by itself provides adequate information. Physiologically, by then we did not know all of the details of the dozens of visual areas in the brain, but we did know enough to be certain that no area contained the panoramic, high-acuity representation of our perceptual experience. The experience had to come from something else, something not coded in a topographic visual map in the brain.

Three solutions to the problem of space constancy have been proposed: the elimination, translation, and evaluation solutions. Physiological and psychophysical evidence allows us to reject all three – no subtraction, compensation, or evaluation need take place. Bridgeman et al. (1994a) offered an alternative "calibration" solution: correct spatiotopic positions are calculated anew from inflow, outflow, and retinal sources for each fixation. There is no need to take previous fixational positions into account; the world appears to be in the same place because nothing tells the visual system that it still is not in the same place. The role of extraretinal signals during saccades, if any, is not to compensate the previous retinal position but to destroy it. Then perception can begin anew with a clean slate during the next fixation interval.

According to a more specific elaboration of this new interpretation, attention shifts to a reference object at the saccade target before a saccade is executed (Deubel et al. 2004). Due to the attention shift, location and visual attributes of the reference object and of surrounding objects are stored in transsaccadic memory. After the saccade, the visual system searches for the reference object within a restricted spatiotemporal "constancy window," which is about 50 msec in duration and is confined to a few degrees around the saccade target. If the object is found, the world is assumed to be stable. Spatial information from the previous fixation is discarded or ignored, and localization proceeds using currently available information. If no other prominent objects are in the region of the saccade landing point, even an object dissimilar to the original saccadic goal object will be accepted as the target if it is in the right position. The positions of other objects in the visual field are then interpreted in terms of the position of the reference object. Only if the object is not found do outflow and other information sources come to bear.

The more radical part of this reanalysis posits that little is carried over from one fixation to the next; we do not build a visual world by pasting together samples calibrated with efference copy, but simply use what is currently available, plus a gist and a few previously attended objects (Irwin, Hollingworth, pers. com.). The stable, rich visual world of our perception is more promise than physiological reality. Extraretinal signals are used in static conditions, though, especially for controlling motor behavior (Bridgeman & Stark 1991).

Evidence for this new position comes from a number of sources, the most dramatic being the demonstrations of change blindness, the inability of observers to identify changes in naturalistic scenes if the change in images is masked by a brief blank of 100 msec or less, a "flicker" paradigm (Simons 1996; Rensink et al. 1997). The interruption need not blank the entire image; if a few "mud splashes" provide visual transients simultaneous with the image change, the change becomes equally invisible (O'Regan et al. 1999). Even the abrupt transient has been shown not to be necessary (Turatto et al. 2003); an image can be ramped down from normal contrast to zero contrast in 1 sec, changed at the instant of zero contrast, and immediately ramped up again, the pattern repeating as in the flicker paradigm. Change blindness is just as strong as in the flicker paradigm, suggesting that it is the diversion of attention rather than abrupt masking transients that underlies the effect.

The importance of change blindness for this article, then, is that a willful inattention to previous images prevents their interfering with present perception. This is the final solution to the space constancy problem.

References

Bell, C. (1823/1974). Idea of a new anatomy of the brain. In P. Cranefield (ed.), *Francois Magendie, Charles Bell and the Course of the Spinal Nerves*. Mt. Kisco, NY: Futura.

Bridgeman, B. (1981). Cognitive factors in subjective stabilization of the visual world. *Acta Psychol* **48**: 111–121.

Bridgeman, B. (1995). Extraretinal signals in visual orientation. In W. Prinz & B. Bridgeman (eds.), *Handbook of Perception and Action Vol. 1: Perception.* London: Academic Press.

Bridgeman, B., Hendry, D., & Stark, L. (1975). Failure to detect displacement of the visual world during saccadic eye movements. *Vision Res* **15**: 719–722.

Bridgeman, B., & Stark, L. (1991). Ocular proprioception and efference copy in registering visual direction. *Vision Res* **31**: 1903–1913.

Bridgeman, B., van der Heijden, A. H. C., & Velichkovsky, B. (1994a). Visual stability and saccadic eye movements. *Behav Brain Sci* **17**: 247–258.

Bridgeman, B., van der Heijden, A. H. C., & Velichkovsky, B. (1994b). How our world remains stable despite disturbing influences. *Behav Brain Sci* **17**: 282–292.

Brune, F., & Lücking, C. (1969). Okulomotorik, Bewegungswahrnehmung und Raumkonstanz der Sehdinge. *Der Nerverarzt* **240**: 692–700.

Deubel, H., Bridgeman, B., & Schneider, W. X. (2004). Different effects of eyelid blinks and target blanking on saccadic suppression of displacement. *Perception & Psychophysics* **66**: 772–778.

Grüsser, O.-J., Krizic, A., & Weiss, L.-R. (1984). Afterimage movement during saccades in the dark. *Vision Res* **27**: 215–226.

Ilg, U., Bridgeman, B., & Hoffman, K.-P. (1989). Influence of mechanical disturbance on oculomotor behavior. *Vision Res* **29**: 545–551.

Leibowitz, H. W., Shupert, C. L., Post, R. B., & Dichgans, J. (1983). Autokinetic drifts and gaze deviation. *Perception & Psychophysics* **33**: 455–459.

Mach, E. (1906). *Die Analyse der Empfindungen und das Verhältnis des Physischen zum Psychischen* (5th ed.).

Mack, A. (1970). An investigation of the relationship between eye and retinal image movement in the perception of movement. *Perception & Psychophysics* **8**: 291–298.

Matin, E. (1974). Saccadic suppression: a review and an analysis. *Psychol Bull* **81**: 899–917.

Matin, L. (1972). Eye movements and perceived visual direction. In D. Jameson & L. Hurvich (eds.), *Handbook of Sensory Physiology* (331–380, vol. **7**, part 3). New York: Springer.

Morgan, C. L. (1978). Constancy of egocentric visual direction. *Perception & Psychophysics* **23**: 61–68.

Nagle, M., Bridgeman, B., & Stark, L. (1980). Voluntary nystagmus, saccadic suppression, and stabilization of the visual world. *Vision Res* **20**: 1195–1198.

O'Regan, J. K., Rensink, R. A., & Clark, J. J. (1999). Change-blindness as a result of "mud-splashes." *Nature* **398**: 34.

Purkinje, J. (1825). Über die Scheinbewegungen, welche im subjectiven Umfang des Gesichtsinnes vorkommen. *Bulletin der naturwessenschaftlichen Sektion der Schlesischen Gesellschaft* **4**: 9–10.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not see: the need for attention to perceive changes in scene. *Psychol Sci* **8**: 368–373.

Simons, D. J. (1996). In sight, out of mind: when object representation fails. *Psychol Sci* **7**: 301–305.

Sperry, R. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *J Comp Physiol Psychol* **43**: 482–489.

Stark, L., & Bridgeman, B. (1983). Role of corollary discharge in space constancy. *Perception & Psychophysics* **34**: 371–380.

Teuber, H.-L. (1960). Perception. In J. Field & H. Magoun (eds.), *Handbook of Physiology*, sect. 1; *Neurophysiology*, vol. **3** (1595–1668). Washington, DC: American Physiological Society.

Turatto, M., Betella, S., Umilta', C., & Bridgeman, B. (2003). Perceptual conditions necessary to induce change blindness. *Visual Cognition* **10**: 233–255.

von Helmholtz, H. (1866). *Handbuch der physiologischen Optik*. Leipzig: Voss.

von Holst, E., & Mittelstaedt, H. (1950). Das Reafferenzprinzip. Wechselwirkungen zwischen Zerntalnervensystem und Peripherie. *Naturwissenschaften* **27**: 464–476.

Wallach, H., & Lewis, C. (1965). The effect of abnormal displacement of the retinal image during eye movements. *Perception & Psychophysics* **81**: 25–29.

8

Intercepting moving objects: do eye movements matter?

ELI BRENNER AND JEROEN B. J. SMEETS

Summary

Due to neuromuscular delays and the inertial properties of the arm people must consider where a moving object will be in the future if they want to intercept it. We previously proposed that people automatically aim ahead of moving objects they are trying to intercept because they pursue such objects with their eyes, and objects that are pursued with the eyes are mislocalized in their direction of motion. To test this hypothesis we examined whether asking subjects to fixate a static point on a moving target's path, rather than allowing them to pursue the target with their eyes, makes them try to intercept the target at a point that the target has already passed. Subjects could not see their hand during the movement and received no feedback about their performance. They did tend to cross the target's path later – with respect to when the target passed that position – when not pursuing the target with their eyes, but the effect of fixation was much smaller than we predicted, even considering that the subjects could not completely refrain from pursuing the moving target as their hand approached it. Moreover, when subjects first started to move, their hands did not aim farther ahead when pursuing the target than when trying to fixate. We conclude that pursuing the target with one's eyes may be important for interception, but not because it gives rise to localization errors that predict the target's displacement during the neuromuscular delay.

8.1 Introduction

It takes tens of milliseconds for visual stimulation of the retina to give rise to activity in the brain (Schmolesky et al. 1998), even longer for neural activity within the brain to result in the contraction of muscles in the arm, and longer yet for the arm to move to its goal. Due to neuromuscular delays and the inertial properties of the arm one must aim ahead of a moving object if one wants to intercept it. How far ahead should depend on how long one expects it to take one's hand to reach the object, and on the object's position and velocity. Surprisingly, making the object appear to move faster by moving the background in the opposite direction does not make people aim farther ahead, and making it appear to move more slowly does not make them aim less far ahead (Smeets & Brenner 1995; Brouwer et al. 2002). To explain this we proposed that the perceived velocity might not be used to predict the point of interception at all (also see Brenner & Smeets 1996). When one wants

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

to intercept a moving object, one follows it with one's eyes. If the position of an object pursued with the eyes were judged to be farther in the direction of ocular pursuit than it really is, then aiming for this position would make one aim ahead of the moving object (Brouwer et al. 2002; Rotman et al. 2005). But why should the perceived position of a moving object be misjudged in such a manner?

Targets that are flashed while subjects are pursuing a moving object with their eyes tend to be mislocalized in the direction of the eye movement (Hazelhoff & Wiersma 1924; Mita et al. 1950; Mitrani et al. 1979; Mateeff et al. 1981). The retinal stimulation by the flash is probably associated with an eye orientation at a later time (Matin et al. 1970; Matin 1986; Schlag & Schlag-Rey 2002). We have proposed that the origin of this misalignment in time is that signals arising from retinal stimulation (some time earlier) are combined with eye movement *command* signals (that will soon give rise to a change in the orientation of the eye) without considering any of the associated neuromuscular delays (Brenner et al. 2001). Pointing tasks reveal a mislocalization that corresponds with an asynchrony of more than 100 msec (Rotman et al. 2004a,b, 2005; Kerzel et al. 2005). This is all for *flashed* targets, but a study showing that the relative positions of flashed and moving targets are not misperceived during pursuit (Nijhawan 2001) suggests that moving targets are mislocalized in a similar manner as flashed ones (although relative positions may be judged independently; Brenner & Cornelissen 2000).

If eye-movement related mislocalization of a moving target is essential for intercepting the target, then we would expect subjects to always hit behind moving targets if they do not move their eyes. But do they? Of course, if they realize that they do so they will compensate for this on subsequent trials. Moreover, the position of the hand may also be misperceived when the eyes are moving, so asking subjects not to pursue the target may not only affect judgments of the target's position, but also that of the hand. To avoid such issues we asked subjects to hit moving targets in an experiment in which they could not see their hand and were not informed about whether they had hit the target. We compared their movements when asked to pursue the moving target with their eyes (which is what they would naturally do) with their movements when asked to fixate a point near where we expected them to intercept the target. If compensating for delays with eye movement–related mislocalization is the only reason for pursuing the target, then except for the hand passing behind the target when fixating, movement trajectories should be similar when fixating and pursuing the target.

8.2 Methods

Thirteen subjects took part in the experiment. Two were the authors. The other eleven were unaware of the hypothesis under study. Each subject took part in one session of 200 trials. There were eight types of trials (twenty-five each) and they were presented in random order. The trials differed in the eye movement that the subject was expected to make (pursuit or fixation), the target velocity (30 or 40 cm/sec), and the position at which we expected the subject to hit the target (left or right). The task was always to hit the target as quickly as

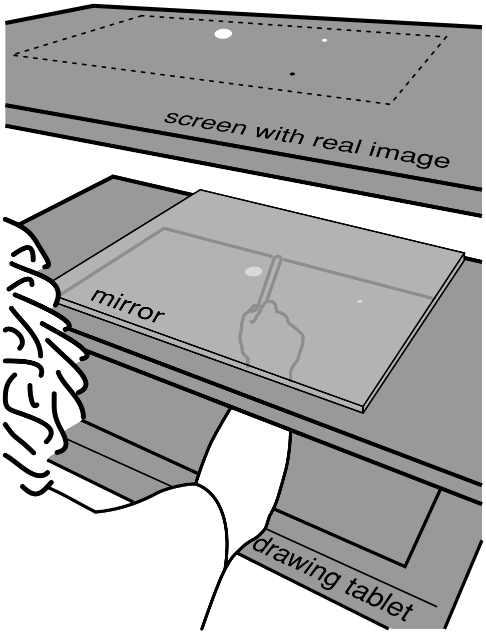


Fig. 8.1 The setup. The white starting point (represented here by the black disk), green target (large white disk), and red fixation point (small white disk) were back-projected from above onto a screen that the subject viewed by way of a mirror. The target always moved rightward. The fixation point was only presented on trials in which the subject was expected to fixate. The image in the mirror was precisely aligned with the surface of a drawing tablet across which the subject moved a pen. The task was to move the tip of the pen *through* the target as quickly as possible. The subject's hand is shown under the mirror to clarify the setup, but the subject could not see his or her hand during the experiment.

possible. To do so subjects moved a "pen" across a large (WACOM A2) drawing tablet. The experiments were conducted in a dimly illuminated room.

Figure 8.1 is a schematic depiction of our setup. The distance between the mirror and the screen (and that between the mirror and the drawing tablet) was 20 cm. The target was a 4-cm-diameter green disk that always moved from left to right. On fixation trials a 1-cm-diameter red disk was visible on the target's path, 8 cm to the right of the center of the drawing tablet. The pen's starting point was a 1-cm-diameter white disk that was 30 cm closer to the subject than the target's path and was also 8 cm to the right of the center of the drawing tablet. Between trials a 1-cm-diameter blue disk indicated where the pen was, to help the subject bring his or her hand to the starting position.

Subjects recognized fixation trials by the fact that the red disk appeared slightly to the right of the center of the tablet. On pursuit trials a similar disk appeared at the position at which the target would later appear, unmistakably to the left of the center of the tablet. This position was not fixed but was determined for each trial on the basis of the time it took the subject to reach the target's path on the previous trial of that condition. By doing so we tried to make the subjects hit the targets near two selected positions, irrespective of their movement times, so that the interception points of trials with different eye-movement instructions would be comparable even if the movement times differed. We used two positions rather than a single position to discourage people from simply always making the same movement. The red disk appeared as soon as subjects held the pen motionless at the starting position. The moving target appeared between 500 and 1000 msec later. On pursuit trials the red disk disappeared when the moving target appeared (at the same position). On fixation trials the red disk remained visible throughout the trial; it occluded part of the target as the target crossed it, but the target always remained visible because it was much larger than the fixation point.

Hand movements (i.e., movements of the tip of the pen) were recorded at 200 Hz. The movements of both eyes were recorded at 250 Hz (Eyelink, SensoMotoric Instruments, Teltow, Germany). The subjects' eyes were about 55 cm from the targets' paths so that the targets' velocities were about 30 and 40 deg/sec. Each session began with a calibration of the eye-movement recordings. To motivate the subjects we gave points for each hit (in inverse proportion to the time it took them to hit the target) and displayed the outcome after the experiment. Because we were looking for small variations in subjects' arm movements we did not want to restrain the subjects in any way. This made it impossible to determine whether the subject's gaze was directed exactly at the moving target or at the fixation point because the subject could move his or her head and body, whereas the Eyelink recorded the orientation of the eye in the head. However, because our hypothesis specifically relates to eye velocity (rather than gaze position) this does not matter. Similar mechanisms to those that we propose could also apply to head and body movements, but the influence that differences in head or body movement speed between the conditions can have on the velocity at which the gaze position changes is presumably negligible in comparison to the influence of differences in eye velocity.

8.3 Results

Fifty-two of the 2600 trials (about 2%) were discarded because the pen did not reach the target's path within 1000 msec of the target appearing or because the data acquisition failed for technical reasons. Tables 8.1 and 8.2 show the average reaction time and movement time for each condition. We first determined the mean value for each subject in each condition and then averaged the thirteen subjects' values for each condition and calculated the standard error of this average. We also subjected the values to a repeated-measures analysis of variance with the eye-movement task, target speed, and position as variables. The reaction time was slightly shorter for the targets that were hit further to the right ($p = 0.002$) and for ones that moved more slowly ($p = 0.046$). The movement time was also slightly shorter for the targets that were hit further to the right ($p = 0.003$). No other differences were significant.

We manipulated the position at which the target appeared to ensure that subjects would hit the targets at about the same position when fixating as during pursuit. This precaution

Table 8.1 *Reaction time in msec (means and standard errors of the thirteen subjects' average values)*

Target Speed	Position	Eyes Fixating Static Point	Eyes Pursuing Target
30 cm/sec	Left	291 (20)	309 (30)
	Right	268 (14)	271 (20)
40 cm/sec	Left	303 (20)	345 (28)
	Right	274 (17)	281 (19)

Table 8.2 *Movement time in msec (means and standard errors of the thirteen subjects' average values)*

Target Speed	Position	Eyes Fixating Static Point	Eyes Pursuing Target
30 cm/sec	Left	369 (35)	379 (31)
	Right	344 (30)	352 (31)
40 cm/sec	Left	368 (35)	360 (30)
	Right	347 (32)	354 (32)

turned out to be superfluous because the reaction and movement times did not depend on the eye movement task. We tried to make subjects hit the targets when they were at two positions (which we refer to as the left and right positions) that were 4 cm apart. In fact, the average distance between the targets at the moment they were hit (which was defined as the moment at which the hand crossed the target's path) was 3.9 cm for the fixation trials and 3.7 cm for the pursuit trials. The overall average target positions were also similar for the two eye-movement tasks (on average the target was 0.2 cm further to the right when it was hit during ocular pursuit).

Although the targets at the left and right positions were almost 4 cm apart (on average) when the hand crossed their paths, the difference between the average positions of the hand was only 2.3 cm when pursuing the target and 1.5 cm when fixating. This can partly be explained by subjects tending to aim toward a similar position as on the previous trial, because on average they hit 0.5 cm further to the right if the previous target was hit on the right ($p = 0.0001$; as evaluated by a similar analysis of variance to that described in the next paragraph, with *previous target position* as an additional variable). The tendency to aim toward a similar position as on the previous trial was slightly stronger (0.6 vs. 0.4 cm) when there was a fixation point, but the difference was not significant ($p = 0.42$). Subjects probably also considered the positions on earlier trials and may also tend to aim toward where they are fixating (moving targets appear to be closer to fixation than they really are; Brenner et al. 2006).

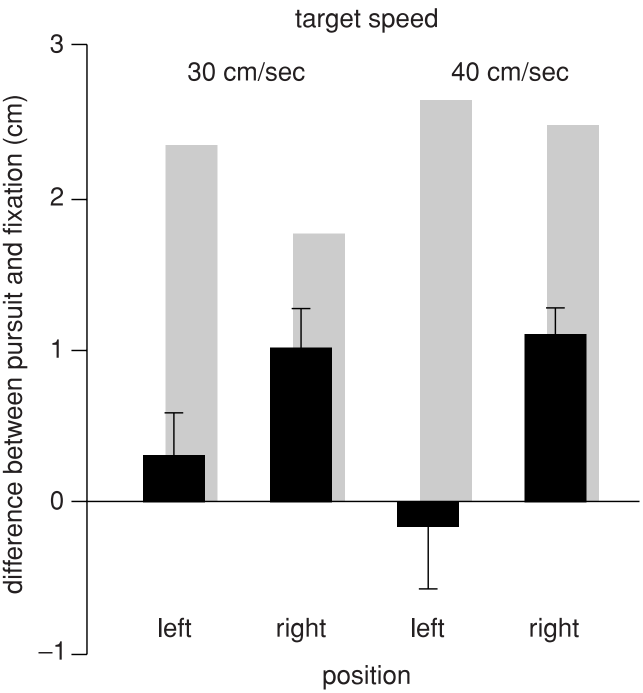


Fig. 8.2 How much further in the target's direction of motion the pen crosses the target's path (relative to the target's position at the time) during ocular pursuit than when trying to fixate (black bars; means and standard errors of the thirteen subjects' average values). The gray bars indicate the minimal extent to which we expected subjects to hit further in the direction of target motion when pursuing the target, considering the eye movements they made.

For a direct evaluation of our proposal we examined the systematic errors that subjects made. How far ahead or behind the target does the pen cross the target's path? There was considerable variability between subjects (average values between –0.2 cm and 5.2 cm) with an overall average value of 2.8 cm (whereby a positive value indicates passing ahead of the target center). A repeated-measures analysis of variance showed that the tendency to hit ahead of the target was larger for targets on the left than on the right (3.8 vs. 1.9 cm; in accordance with the tendency mentioned in the previous paragraph; $p < 0.0001$) and was slightly larger for the faster targets (3.0 vs. 2.6 cm; $p = 0.003$). Most importantly, it was slightly larger during pursuit than during fixation (3.1 vs. 2.5 cm; $p = 0.03$), especially on the right ($p = 0.0009$ for the interaction between position and eye-movement task). The extent to which subjects hit further ahead during pursuit is shown in Fig. 8.2.

Figure 8.3 shows the average pen-movement paths for the two eye-movement instructions. These paths were constructed by resampling the lateral position of the pen for fifty equidistant sagittal steps from the pen's starting position to the target's path (using linear interpolation). The resampled positions for individual movements were then first averaged across trials for each subject and condition, and then across subjects. The average paths

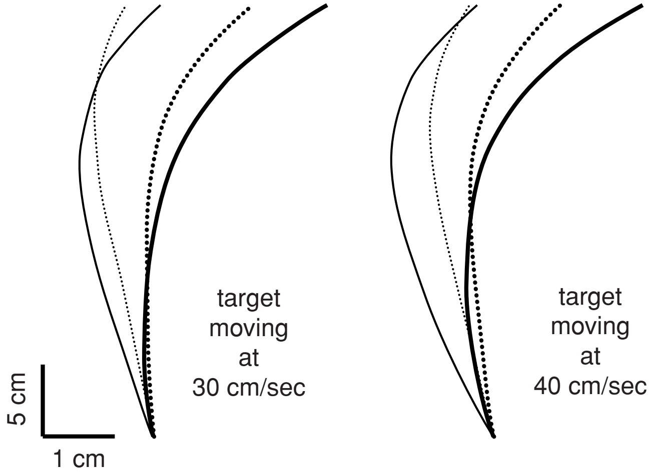


Fig. 8.3 The average path of the hand (pen) in each of the eight conditions. Thick curves are for the target that was to be hit on the right and thin ones for the target that was to be hit on the left. Solid curves are for movements while the subject pursued the target with his or her eyes. Dotted curves are for movements while the subject tried to fixate a static disk near where he or she was expected to hit the target. The lateral movement has been exaggerated for clarity (see different scales).

are slightly curved. This curvature differs systematically between the two eye-movement tasks. Contrary to our prediction, the subjects' hands initially appear to have been heading less far ahead of the target when the subject was pursuing the target with his or her eyes, rather than further ahead. This too may be caused by a tendency to underestimate the retinal eccentricity of the target (judging it to be closer to where one is fixating).

Figure 8.4 shows the average velocity of the eye and hand near the moment that the hand passed the target's path. The orientations of the two eyes were first averaged and smoothed with a Gaussian curve with a standard deviation of 8 msec. Velocities of eye and hand were determined by dividing the distance between consecutive samples by the time interval between them. These values were attributed to the moment between the two samples. The velocities were then averaged across trials for each subject and condition. The mean velocity profile (curve) and standard error across subjects (shaded area) are shown for each condition. When the task was to pursue the target with their eyes (solid curves), the subjects' eyes moved at about the same velocity as the target (indicated by the dashed horizontal lines) during the last 200 msec before the hand passed the target's path (top panels). When the task was to fixate (dotted curves), the eye gradually accelerated as the target and hand approached the fixation point. The hand decelerated in the sagittal direction as it approached the target's path (bottom panels of Fig. 8.4) and accelerated to the right (central panels). The hand moved faster to the right near the time of the hit when the target was on the right (thick curves) and when the subjects were following the target with their eyes (solid curves).

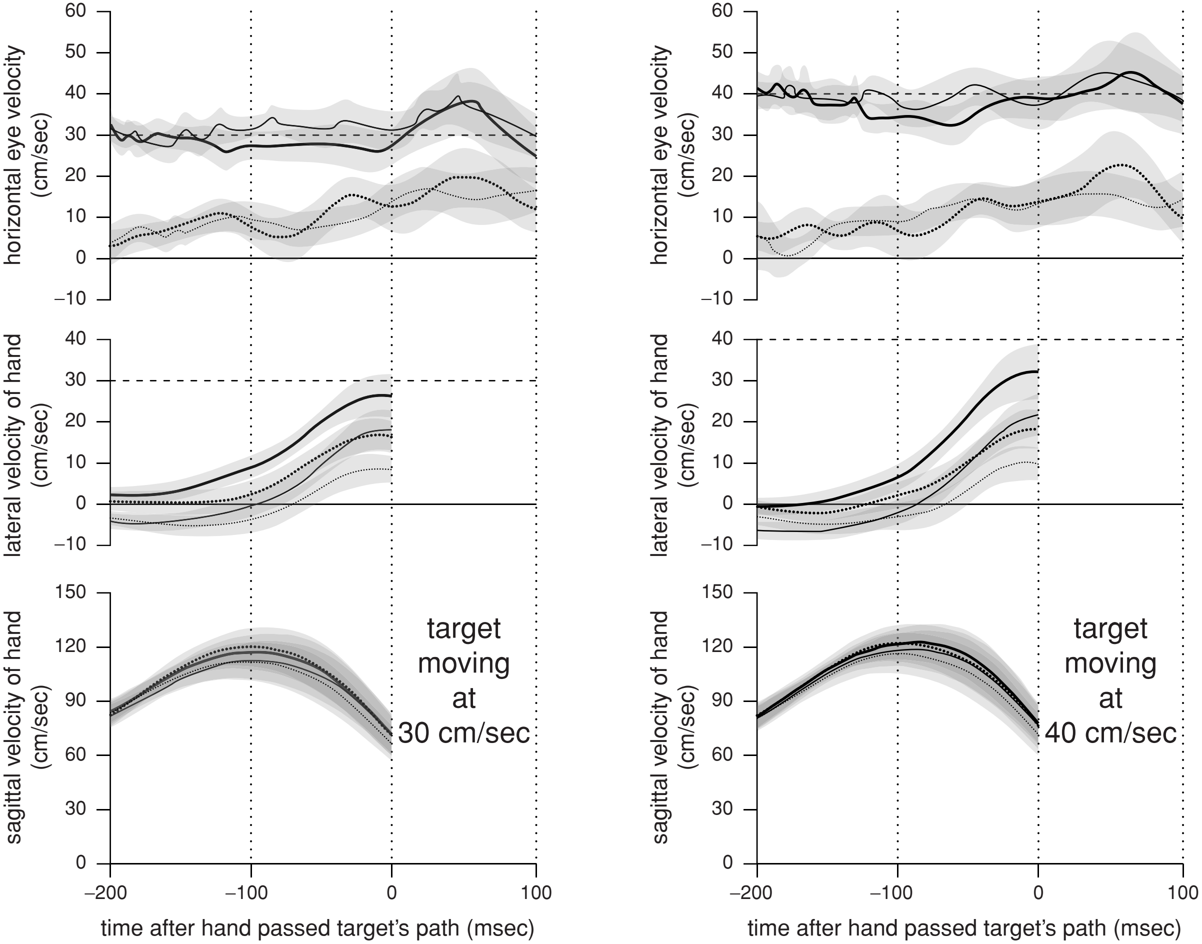


Fig. 8.4 The average velocity of the eye and hand (pen) near the moment that the subjects tried to intercept the target. Thick and thin curves are for movements in which we expected the target to be hit on the right and left, respectively. Solid curves are for movements while the subject tried to pursue the target with his or her eyes. Dotted curves are for movements while the subject tried to fixate a static disk near where the target was to be hit. The shaded areas indicate the standard errors across subjects.

The critical question was whether subjects would hit further back along the target's path if they were forced to fixate while hitting. We initially expected them to hit at least 3 or 4 cm behind the target when forced to fixate. These are the distances that the targets move during 100 msec, and we anticipated that pursuing the targets gives rise to a localization error that corresponds with the targets' displacement during at least 100 msec. However, because subjects also pursued the targets with their eyes to some extent when trying not to do so, we have to refine our expectation. Our refined estimate of the expected effect depends on the difference between the extent to which subjects pursue the target on pursuit and fixation trials, which changes during the movement. The minimal difference that could be considered consistent with the proposal that mislocalizing targets during pursuit overcomes neuromuscular delays is 100 msec times the difference in the velocity of pursuit at the moment that the hand passed the target's path. This is the prediction shown by the gray bars in Fig. 8.2. It is clear that the tendency to hit further along the target's path when pursuing the target with the eyes than when fixating is much weaker than we predicted. Moreover, the extent to which individual subjects hit further ahead on pursuit trials than on fixation trials was not significantly correlated with the extent to which their average eye-movement velocity differed between the two kinds of trials.

On average, subjects crossed the target's path further ahead of the target during pursuit than during fixation, but the extent to which they did so was only 25% of the predicted effect. Because our prediction did not distinguish between pursuit and saccades, we averaged the eye movements irrespective of whether saccades were made or not. To make sure that our conclusion would not have been different if we had restricted ourselves to *smooth* pursuit, we also selected the trials in which there was certainly no saccade near the critical moment: when the velocity of the eye did not exceed 100 deg/sec during the last 100 msec before the target was intercepted. For those trials (92% of fixation trials and 89% of pursuit trials), the extent to which subjects hit further ahead of the target during pursuit was 30% of the predicted effect. This is slightly more than we found when we included all trials. The increase is mainly caused by the fact that the *predicted* effect is smaller (on average by 0.16 cm) if trials with saccades are excluded, which is not too surprising because 85% of the saccades during pursuit trials were in the direction in which the target was moving, whereas only 51% of the saccades in the fixation trials were in the direction of target motion (for the left position on fixation trials most saccades were to the left, opposite the direction of target motion). The tendency to hit further ahead of the target during pursuit was (on average) 0.065 cm larger on trials without saccades. Although these differences may be interesting, it is clear that excluding trials with saccades does not change our main conclusions.

8.4 Discussion

Apart from disproving our hypothesis about the relationship between manual interception and the mislocalization of targets flashed during pursuit, this study reveals two interesting phenomena. The first is that the hand's path toward the target depends on the eye movements the subject makes (Fig. 8.3). The hand's sagittal velocity hardly differs between the conditions (Fig. 8.4). It decreases as the hand approaches the target, but that is probably just the result of the hand reaching full extension and the end of the drawing tablet. The acceleration to the right, however, does differ between the conditions (Fig. 8.4). In all cases the hand curves slightly to move along with the target near the moment of interception (perhaps to reduce the influence of errors in judging one's own sagittal movement; Brenner & Smeets 2005), but when there was a static fixation point the path was less curved than when the eyes pursued the target (Fig. 8.3).

The difference between the paths could be explained by subjects underestimating the moving target's retinal eccentricity (Brenner et al. 2006) because if so subjects will initially aim closer to the fixation point when not pursuing the target and, because they are fixating a point that the target is still to pass, this means they will initially aim further ahead of the target. If the retinal eccentricity of the moving target is systematically underestimated, then the target's apparent position will change more rapidly during pursuit than when fixating. This could explain why the hand ended up moving faster to the right during pursuit (Fig. 8.4), and perhaps even why the final position was slightly further ahead of the target in that case (Fig. 8.2). Thus the different curvatures of the paths, different final lateral velocities of the hand, and different positions at which the hand crossed the target's path may all result from underestimating the retinal eccentricity of the moving target. Alternatively, subjects may tend to move their hand toward where they are looking for some other reason, or they may just underestimate the target's velocity during pursuit (e.g., Dichgans et al. 1975; Sumnall et al. 2003), either of which could also account for the difference between the paths.

The second interesting phenomenon is that subjects were unable to maintain fixation near the moment they hit the target. This is not simply a reflexive response of the eye to the target crossing the fixation point because the eye did not speed up earlier or more strongly when the target was to be hit on the right, although such targets crossed the fixation point 133 or 100 msec earlier (for targets moving at 30 and 40 cm/sec, respectively). It also cannot be considered proof that the eye and hand are functionally linked during interception, supporting the many examples of failures to independently move the eye and hand (e.g., Lunenburger et al. 2000; Neggers & Bekkering 2001, 2002; Horstmann & Hoffmann 2005), because subjects also fail to fixate when performing tasks that require that one attends to a moving object without moving one's hand toward it (Khurana & Kowler 1987). It is interesting that the eye even seems to rotate to the right, in the same direction as the target and the hand, when the target is hit to the left of the fixation point (thin dashed lines in top panel of Fig. 8.4; note that we did not remove the occasional leftward saccades when calculating these average paths). For a better understanding of these eye movements, we need to have more complete information about the direction of gaze.

The relationship between the eye and hand movements was not as we predicted. Even considering the failure to fixate, subjects did not hit even nearly as much further ahead of the targets during pursuit as we had predicted. The predictions shown in Fig. 8.4 are low estimates for the expected difference because the mislocalization during pursuit usually corresponds to more than 100 msec of target motion, and the difference in eye velocity would be larger if we were to consider an earlier moment before passing the target's path (remember that our proposal applies to the whole movement). Moreover, the movement paths curved differently when fixating than when following the target with the eyes, which would not be so if subjects only consistently misjudged the target's position in its direction of motion. Thus we conclude that eye movements play a role in interception, but that this role is not mediated by the perceptual mislocalization that is revealed when targets are flashed during eye movements.

References

Brenner, E., & Cornelissen, F. W. (2000). Separate simultaneous processing of egocentric and relative positions. *Vision Res* **40**: 2557–2563.

Brenner, E., & Smeets, J. B. J. (1996). Hitting moving targets: co-operative control of 'when' and 'where.' *Hum Mov Sci* **15**: 39–53.

Brenner, E., & Smeets, J. B. J. (2005). Intercepting moving targets: why the hand's path depends on the target's velocity. In B. E. Rogowitz, T. N. Pappas, & S. J. Daly (eds.), *Human Vision and Electronic Imaging X*, *Proc. SPIE-IS&T Electronic Imaging* (374–384). SPIE Vol. **5666**.

Brenner, E., Smeets, J. B. J., & van den Berg, A. V. (2001). Smooth eye movements and spatial localisation. *Vision Res* **41**: 2253–2259.

Brenner, E., van Beers, R. J., Rotman, G., & Smeets, J. B. J. (2006). The role of uncertainty in the systematic spatial mislocalization of moving objects. *J Exp Psychol Hum Percept Perform* **32**: 811–825.

Brouwer, A., Brenner, E., & Smeets, J. B. J. (2002). Hitting moving objects: is target speed used in guiding the hand? *Experimental Brain Research* **143**: 198–211.

Dichgans, J., Wist, E., Diener, H. C., & Brandt, T. (1975). The Aubert-Fleischl phenomenon: a temporal frequency effect on perceived velocity in afferent motion perception. *Experimental Brain Research* **23**: 529–533.

Hazelhoff, F. F., & Wiersma, H. (1924). Die Wahrnehmungszeit. *Zeitschrift für Psychology* **96**: 171–188.

Horstmann, A., & Hoffmann, K. P. (2005). Target selection in eye-hand coordination: do we reach to where we look or do we look to where we reach? *Experimental Brain Research* **167**: 187–195.

Kerzel, D., Aivar, M. P., Ziegler, N. E., & Brenner, E. (2005). Mislocalization of flashes during smooth pursuit hardly depends on the lighting conditions. *Vision Res* **46**: 1145–1154.

Khurana, B., & Kowler, E. (1987). Shared attentional control of smooth eye movement and perception. *Vision Res* **27**: 1603–1618.

Lunenburger, L., Kutz, D. F., & Hoffmann, K. P. (2000). Influence of arm movements on saccades in humans. *Eur J Neurosci* **12**: 4107–4116.

Mateeff, S., Yakimoff, N., & Dimitrov, G. (1981). Localization of brief visual stimuli during pursuit eye movements. *Acta Psychologica* **48**: 133–140.

Matin, L. (1986). Visual localization and eye movements. In K. R. Boff, L. Kaufman, & J. P. Thomas (eds.), *Handbook of Perception and Human Performance; Volume 1; Sensory Processes and Perception* (1–40). New York: Wiley-Interscience.

Matin, L., Matin, E., & Pola, J. (1970). Visual perception of direction when voluntary saccades occur: II. Relation of visual direction of a fixation target extinguished before a saccade to a subsequent test flash presented before the saccade. *Perception & Psychophysics* **8**: 9–14.

Mita, T., Hironaka, K., & Koike, I. (1950). The influence of retinal adaptation and location on the "Empfindungszeit." *The Tohoku Journal of Experimental Medicine* **52**: 397–405.

Mitrani, L., Dimitrov, G., Yakimoff, N., & Mateeff, S. (1979). Oculomotor and perceptual localization during smooth eye movements. *Vision Res* **19**: 609–612.

Neggers, S. F., & Bekkering, H. (2001). Gaze anchoring to a pointing target is present during the entire pointing movement and is driven by a non-visual signal. *J Neurophysiol* **86**: 961–970.

Neggers, S. F., & Bekkering, H. (2002). Coordinated control of eye and hand movements in dynamic reaching. *Hum Mov Sci* **21**: 349–376.

Nijhawan, R. (2001). The flash-lag phenomenon: object motion and eye movements. *Perception* **30**: 263–282.

Rotman, G., Brenner, E., & Smeets, J. B. J. (2004a). Mislocalization of targets flashed during smooth pursuit depends on the change in gaze direction after the flash. *J Vision* **4**: 564–574.

Rotman, G., Brenner, E., & Smeets, J. B. J. (2004b). Quickly tapping targets that are flashed during smooth pursuit reveals perceptual mislocalisations. *Experimental Brain Research* **156**: 409–414.

Rotman, G., Brenner, E., & Smeets, J. B. J. (2005). Flashes are localised as if they were moving with the eyes. *Vision Res* **45**: 355–364.

Schlag, J., & Schlag-Rey, M. (2002). Through the eye, slowly: delays and localization errors in the visual system. *Nat Reviews Neurosci* **3**: 191–200.

Schmolesky, M. T., Wang, Y., Hanes, D. P., Thompson, K. G., Leutgeb, S., Schall, J. D., et al. (1998). Signal timing across the macaque visual system. *J Neurophysiol* **79**: 3272–3278.

Sumnall, J. H., Freeman, T. C., & Snowden, R. J. (2003). Optokinetic potential and the perception of head-centred speed. *Vision Res* **43**: 1709–1718.

9

The utility of visual motion for goal-directed reaching

DAVID WHITNEY, IKUYA MURAKAMI, AND HIROAKI GOMI

Summary

Visual information is crucial for goal-directed reaching. Recently a number of studies have shown that motion in particular is an important source of information for the visuomotor system. For example, when reaching for a stationary object, nearby visual movement, even when irrelevant to the object or task, can influence the trajectory of the hand. Although it is clear that various kinds of visual motion can influence goal-directed reaching movements, it is less clear how or why they do so. In this chapter, we consider whether the influence of motion on reaching is unique compared to its influence on other forms of visually guided behavior. We also address how motion is coded by the visuomotor system and whether there is one motion processing system that underlies both perception and visually guided reaching. Ultimately, visual motion may operate on a number of levels, influencing goal-directed reaching through more than one mechanism, some of which may actually be beneficial for accurate behavior.

9.1 Introduction

Visual motion is constantly produced as we move our eyes and head and as objects move in the world. The visuomotor system, therefore, faces a serious challenge in that it must register target as well as background motion and then segment these different sources of motion in order to direct actions to objects. Over the last three decades, a broad and expanding literature has examined how the visuomotor system processes and uses visual motion in goal-directed behavior. Although there is still a great deal that remains unknown, a coherent picture is beginning to emerge, and the results of recent studies have overturned traditional intuitions about what sorts of visual motion information are used to guide action.

9.1.1 Moving objects

The use of visual motion by the visuomotor system has been studied in a number of different situations, including that from moving objects, entire scenes, or just backgrounds (Fig. 9.1). Of these, the visual motion of objects has been studied most extensively. This is not surprising, as some of the most impressive human motoric abilities involve goal-directed behaviors to moving objects on brief time scales. Sports such as baseball, cricket,

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

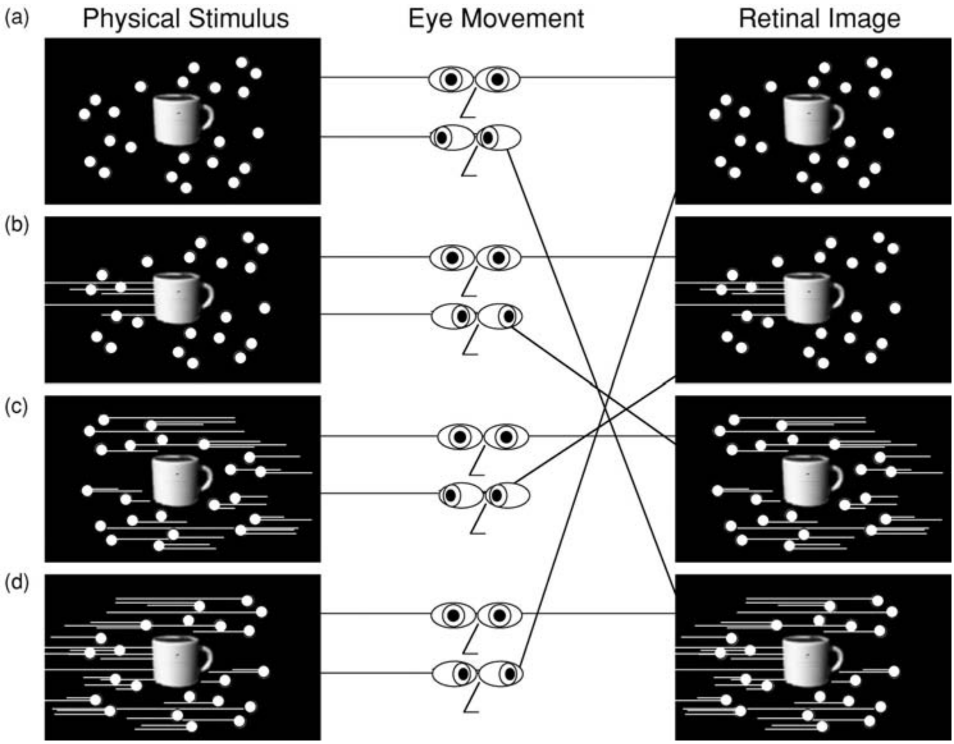


Fig. 9.1 Several types of retinal motion (right panel) produced by various combinations of physical stimuli (left panel) and eye movements (or head or body movements, middle panel). The mug represents a target object and the white dots a structured background. (a) physically static world can create either a static retinal image or one that moves opposite the direction of an eye movement. (b) A moving object with a static background can produce the same retinal image (with no ego motion) or retinal motion of the background alone (with fixation of the target). (c) The movement of the background alone (e.g., fixating the stereo as we drive in a car) can produce background motion with a retinally static target. (d) The movement of targets with the background can produce coherent static or moving retinal images. Because different combinations of physical stimuli and eye movements can produce identical patterns of retinal motion, the visuomotor system needs to segment target motion from motion of the background.

jai-alai, and others present ample evidence that the human brain is capable of rapidly coding object position and motion (Land & McLeod 2000; Gray 2002) and that there are predictive mechanisms at work (Savelsbergh et al. 1991; Tresilian 1993; Regan 1997). The motion complex MT+ (the human homologue of monkey MT and MST) appears to play an important role in these abilities (Schenk et al. 2005).

9.1.2 Moving backgrounds

Although moving objects provide a means of testing the limits of the visuomotor system, motion of individual target objects is not the only type of motion we face. More commonly, we face situations in which the target and background move together (Fig. 9.1(d)) or the target is stationary while the background moves (Fig. 9.1(c)). This is because the world (often including the target object) is usually stationary while we are moving. This not only causes retinal motion of the target but retinal motion of the entire scene. Similarly, even when the target object is moving in the world, we often move our eyes and head to track the object. In both instances, there is background retinal motion that the visuomotor system must explicitly distinguish from the target object (and then either ignore or use it, if it is informative).

9.1.3 Benefits of visual motion: optic flow

One potentially useful type of background retinal motion information extensively studied is optic flow. As we move around the world, characteristic patterns of retinal motion are produced that may be informative about our body movements (Lee 1980; Gibson 1986). For example, moving forward through a stationary scene produces expanding optic flow. Many species of animals, including humans, use this kind of optic flow to gauge and adjust posture (Lee & Aronson 1974; Lee & Reddish 1981; van Asten et al. 1988; Previc 1992; Warren et al. 2001). It has been shown that neurons in visual area MST respond to different kinds of optic flow (e.g., expanding, contracting, shearing) (Saito et al. 1986; K. Hikosaka et al. 1988; Tanaka & Saito 1989; Duffy & Wurtz 1991a,b). Physiological experiments in monkeys have demonstrated that stimulation of neurons showing optical flow preferences (in MST) influences judgments of heading (Britten & van Wezel 1998). Consistent with the monkey physiology, neuroimaging studies in humans indicate that there is a homologue of the MST area that exhibits preferential activation to patterns of optic flow (Greenlee 2000; Dukelow et al. 2001; Huk et al. 2002). However, many other high-level motion areas that respond to complex visual motion, such as biological motion and inferred motion (Culham et al. 2001), do not show the selectivity for optic flow.

Optic flow is not simply a property of backgrounds or scenes, but may be used for visually guided behaviors toward or away from individual objects. One possible by-product of optic flow computations is time-to-collision (τ; Lee & Reddish 1981). As an object looms or approaches our eyes, the expanding retinal image (optic flow) may afford information about the time-to-collision. Physiological evidence in pigeons (Wang & Frost 1992) as well as behavioral evidence in humans and diving gannets suggests that the visual system may use a derivative of optic flow information to estimate time-to-collision (Savelsbergh et al. 1991; Regan 1997), although this is controversial (Tresilian 1993).

9.1.4 Visually guided reaching with background motion

The evidence that visual motion is used to guide behavior so far comes from studies of posture and timing (e.g., τ). Recently, however, it has been found that background retinal motion can influence the direction of reaching movements as well. For example, when subjects reach a stationary target, a moving background can cause shifts in the trajectory of the reaching movement and endpoint consistent with the direction of the background motion (Mohrmann-Lendla & Fleischer 1991; Brenner & Smeets 1997; Yamagishi et al. 2001; Whitney et al. 2003; Saijo et al. 2005).

The effect of background motion on reaching is distinct from the effect of optic flow on posture. Most forms of optic flow (especially expanding and contracting flow fields) are perfectly predictive of head movement and largely predictive of whole-body movements. It is therefore intuitive that this sort of visual information could be used to control posture, locomotion, or body position. On the other hand, the influence of background motion on reaching is counterintuitive because it is neither predictive of hand position or motion, nor is it predictive of the relationship between the hand and the target. Several authors have proposed explanations for the influence of background motion on reaching, but the issue remains controversial. In this chapter we review the evidence for the influence of visual motion on reaching, consider alternative underlying mechanisms, and discuss whether the effects are a beneficial adaptation.

9.2 Object motion and the double-step paradigm

When reaching to a moving object, the visuomotor system could predict future hand position relative to the current hand position by using a forward model (Desmurget & Grafton 2000). This could allow for rapid corrections in response to either target or effector (hand or mouse) displacements and may help reduce the error introduced by sensory and motor delays. An important question over the last decade has been to evaluate what sorts of visual information are used for online control of reaching.

Many studies investigating goal-directed reaching to moving targets demonstrate that the visuomotor system monitors target motion (or perhaps relative position or motion between the hand and moving object) and uses this information to guide the hand. Comparable results have been obtained whether the target physically jumps (Goodale et al. 1986; Pelisson et al. 1986; van Sonderen et al. 1988, 1989; Paulignan et al. 1991b; Prablanc & Martin 1992; Day & Lyon 2000; Brenner & Smeets 2003) or moves continuously (Brenner et al. 1998; Schenk et al. 2004). Subjects are able to make rapid corrections to ongoing behavior.

One of the strongest examples of the speed with which the visuomotor system can respond to changing visual information is the perturbation or double-step paradigm (Bridgeman et al. 1979). In this technique, a target is presented to which subjects reach. During the reach, the target can be displaced in position. The reaction of the hand to the displaced target occurs within 100–150 msec (Goodale et al. 1986; Pelisson et al. 1986; van Sonderen et al. 1988, 1989; Castiello et al. 1991; Paulignan et al. 1991b; Prablanc & Martin 1992; Brenner & Smeets 1997). Even if the displaced target is presented during a saccade, preventing awareness of its displacement, subjects can still make fast and accurate corrections (Bridgeman et al. 1979; Goodale et al. 1986; Pelisson et al. 1986). Although very different measurements of the visuomotor reaction to perturbed targets have been used (including derivatives of hand position, such as velocity and acceleration), the results have been uniformly similar. The brevity of the reaction time – the visual feedback delay – shows that visual information is important in the online control of action even late into the movement (Desmurget & Grafton 2000).

The double-step paradigm has provided a rich source of information about how the visuomotor system programs and executes action. For example, it has shown that the fast visual feedback necessary for motor control is not dependent upon awareness of the target's displacement or motion (Goodale et al. 1986; Pelisson et al. 1986). This points to a direct route independent of conscious vision and provides support for, at least, the moderate form (Jacob & Jeannerod 2003) of the dual visual systems hypothesis (Trevarthen 1968; Schneider 1969; Ingle 1973; Goodale & Milner 1992). Double-step studies have also demonstrated that reaching movements can be guided and modified online, whether or not subjects fixate on a static point (Goodale et al. 1986; Pelisson et al. 1986; van Sonderen et al. 1988, 1989; Castiello et al. 1991; Paulignan et al. 1991b; Prablanc & Martin 1992; Brenner & Smeets 1997; Diedrichsen et al. 2004). This is important for evaluating the hypothesis that the hand is guided or anchored by the direction of gaze (Paillard 1982; Binsted et al. 2001; Neggers & Bekkering 2001). Although it is generally true that we reach where we are looking (Ballard et al. 1992; Pelz et al. 2001; Soechting et al. 2001), this is not exclusively the case, and bimanual tasks (particularly bimanual double-step tasks) provide adequate evidence for not needing to fixate the endpoint goal of the hand (Diedrichsen et al. 2004). The double-step paradigm has provided even stronger evidence that corrections to fast-reaching movements can occur on an extremely brief time scale, even without changes in eye position and based only on peripheral retinal information (Paillard 1996). Therefore, although eye position certainly contributes to visuomotor control, and foveal location or gaze direction may be an important source of information with which to guide the hand, it is only one of many sources of information that contribute to visually guided behavior.

What remains unclear from the double-step studies is whether the target position or motion (or both) are used for online guidance of the hand. One might ask whether a distinction should even be drawn between the motion and position of an object, because these properties are necessarily *physically* interdependent. The visual system, however, need not code object motion and position as a single dimension. Along with spatiotemporal sensitive mechanisms that simultaneously code the position and motion of an object (Burr et al. 1986; Geisler 1999; Burr & Ross 2002; Nishida 2004), several visual illusions suggest that the visual system may also have independent mechanisms for coding object motion and position (Whitney 2002). Therefore, it becomes meaningful to ask whether rapid corrections to visually guided reaching are driven by object motion, shifts in the coded locations of targets due to the motion, or a combination of these. Unfortunately, to date this question has not been fully addressed (see Section 5).

9.3 Background scene motion

The above experiments have been extended by showing that both reaching and pointing movements tend to deviate in the direction of background retinal motion (Mohrmann-Lendla & Fleischer 1991; Brenner & Smeets 1997; Proteau & Masson 1997; Yamagishi et al. 2001; Whitney et al. 2003; Saijo et al. 2005), not just target motion (Smeets & Brenner 1995a). There are several surprising characteristics of this effect. First, even in the presence of a stationary target, background retinal motion influences the reach (Brenner & Smeets 1997; Whitney et al. 2003; Saijo et al. 2005). Second, retinal motion can be separated by many degrees from the target and still be influential (Whitney et al. 2003; Saijo et al. 2005). Third, reaction time to background retinal motion is around 120–150 msec (Brenner & Smeets 1997; Whitney et al. 2003; Saijo et al. 2005), which is just as fast as reactions to changes in target location or motion of the target itself in double-step studies (Prablanc & Martin 1992). Visual motion during either the programming (prior to the reaction) or execution influences the trajectory of the reach (Whitney et al. 2003). Finally, the percept of the target (its motion or position) can be dissociated from the hand's trajectory, demonstrating that the hand does not simply follow perceived target position or motion (Yamagishi et al. 2001; Whitney et al. 2003). Rather, the visuomotor system uses visual motion to directly control the hand.

9.3.1 Underlying cause of background motion's influence on reaching

There is not consensus about the underlying cause of the influence of visual motion of the background on the trajectory and endpoint of goal-directed reaching movements. It has been argued that background visual motion shifts the coded locations of objects (Whitney et al. 2003). This is an "indirect" mechanism, because the visuomotor response is modulated by the target position (or shifted position). Alternatively, it has been suggested that background motion triggers a passive manual following response (Saijo et al. 2005) akin to the ocular following response (Kawano & Miles 1986; Miles et al. 1986; Kawano et al. 1994; Masson et al. 2001). This is a "direct" effect of motion on reaching, as the representation of the target position is not shifted.

To distinguish between direct and indirect influences of visual motion on visuomotor control, it is helpful to consider a distinction made in the visual motion perception literature between velocity- and position-based motion computations. Visual motion could be an inference based on perceiving the position of an object at two time points (Ullman 1979; Nakayama & Tyler 1981; Cavanagh 1992; Del Viva & Morrone 1998; Seiffert & Cavanagh 1998; Derrington et al. 2004). There is evidence for this sort of motion detector (Cavanagh 1992; Lu & Sperling 1995), which may involve attention and could be helpful in tracking camouflaged or second-order (feature or contrast-defined) motion. This mechanism indirectly measures motion and cannot recover it without a means of coding object position. The more familiar, dominant, and common sort of motion detector is sensitive to velocity and is considered a "direct" measure of motion because it does not require multiple samples over space or time (Adelson & Bergen 1985; van Santen & Sperling 1985; Watson & Ahumada 1985). Even when an object position has not been identified or tracked, a velocity detector can still register motion.

The distinction between directly and indirectly sensed visual motion carries over to visuomotor control. When programming and executing reaching movements, the visuomotor system could, independent of any target representation, monitor visual motion information and use this to guide the hand (Saijo et al. 2005; Gomi et al. 2006). Alternatively, visual motion could influence the coded locations of targets, and it is these target positions that are monitored throughout the execution of the reach (Whitney et al. 2003). Cases of akinetopsia (Zihl et al. 1983) show that patients can perceive object position but not object motion, whereas cases of blindsight (Cowey & Stoerig 1991) reveal an unawareness of object position coupled with intact perception of nonlocalized motion. Although blindsight and akinetopsia are somewhat controversial, the possibility that motion directly and/or indirectly contributes visuomotor control is plausible. Evidence for each possibility is reviewed in the following.

9.3.2 Shifted target positions

One explanation for the deviations in the trajectory and endpoints of the reach is that the target position that serves to guide the hand is shifted by background motion (Brenner & Smeets 1997; Whitney et al. 2003). It is well known that visual motion can influence the perceived locations of both static and moving objects (Ramachandran & Anstis 1990; De Valois & De Valois 1991; Nishida & Johnston 1999; Whitaker et al. 1999; Hayes 2000; Whitney & Cavanagh 2000; McGraw et al. 2002) (see Fig. 9.2 and Whitney 2002, for a review). For example, the briefly flashed objects in Fig. 9.2(c) appear shifted in the direction of the nearest visual motion. Is it possible that visual motion shifts either the perceived location or the represented location of targets, thus indirectly causing a deviation in the hand's trajectory? There are three possibilities: the reach may simply follow the perceived shift in target location; the reach may follow a shift in the coded location of the target independent of what is perceived; or the reach may not depend on shifts in the perceived or coded target location.

9.3.2.1 Dissociating perception and action

Is the influence of visual motion on reaching due to a visual illusion, or does the visuomotor system rely on visual motion information that bypasses conscious vision? Yamagishi et al. (2001) found that pointing movements to a drifting Gabor pattern were differentially shifted relative to the perceived location of the Gabor (also see Ashida 2004). Although the directionality of the illusion was comparable for perception and action, the magnitudes differed (but c.f., Kerzel & Gegenfurtner 2003, 2005) for counterevidence). Morhmann-Lendla & Fleischer (1991) and Brenner & Smeets (1997) obtained similar results in that the trajectories of reaching movements directed to physically static objects deviated slightly in the direction of superimposed moving backgrounds. The perceived position of continuously visible static targets in this situation, however, is generally veridical, while displaying, if anything, induced motion or motion contrast (Brenner & Smeets 1994; Smeets & Brenner 1995b). Similar dissociations between perception and action have been reported in studies of frame of reference, induced motion, and induced Roelofs effects, although their relation to the influence of background motion on reaching is unclear (Bridgeman et al. 1979, 1981,

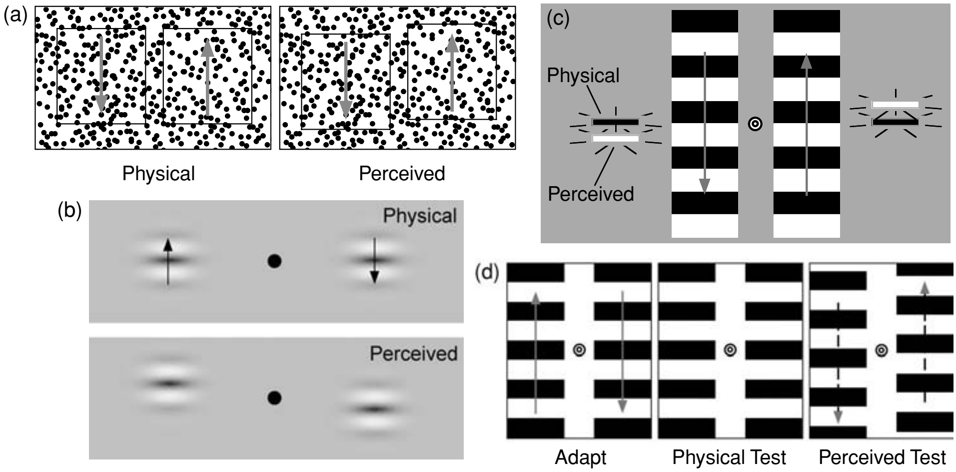


Fig. 9.2 The influence of visual motion on perceived position. (a) Two kinetically defined squares contain dots drifting in opposite directions. The black lines were not visible; the squares were defined only by the relative motion of the drifting dots. The right panel shows the perceived positions of the motion-defined squares, which are displaced in the direction of the motion. (b) Drifting a luminance-modulated carrier (sine wave) inside a static Gaussian contrast envelope causes the static envelope to appear shifted in position, in the direction of the visual motion. The magnitude of the effects in (a) and (b) peak when the luminance of the drifting regions is equated to that of their surroundings. (c) Motion in one region influences the perceived positions of briefly flashed static objects separated from the motion. (d) The motion aftereffect, following motion adaptation, can be accompanied by a concurrent shift in the perceived position of the test pattern. Figure adapted from Whitney (2002) with permission.

1997; Sheth & Shimojo 2000; Dassonville et al. 2004; Post & Welch 2004). Nonetheless, a variety of paradigms suggest that the influence of visual motion on reaching is not simply based on misperceiving the target motion or position.

9.3.2.2 Relative speed of perception and action

In order to gauge the relative time course of motion's influence on visuomotor behavior and perception, Whitney et al. (2003) employed a motion-reversal paradigm that involved presenting a moving pattern for an unpredictable period of time followed by a reversal in its motion. A brief static target was flashed several degrees away from the visual motion at an unpredictable moment – either before, simultaneous with, or after the motion reversal (Fig. 9.3). When the target was presented during continuous motion (well before the reversal), the hand deviated in the direction of the background motion (similar to Mohrmann-Lendla & Fleischer 1991; Brenner & Smeets 1997; Yamagishi et al. 2001). However, when the target was presented just before the motion reversal (circled points in Fig. 9.3(b)), the hand initially deviated in the direction of the initial visual motion and then abruptly reversed the direction of its deviation (Fig. 9.3(c)). In other words, the trajectory of the hand closely mimicked the motion reversal – albeit delayed in time. Visual motion therefore influences

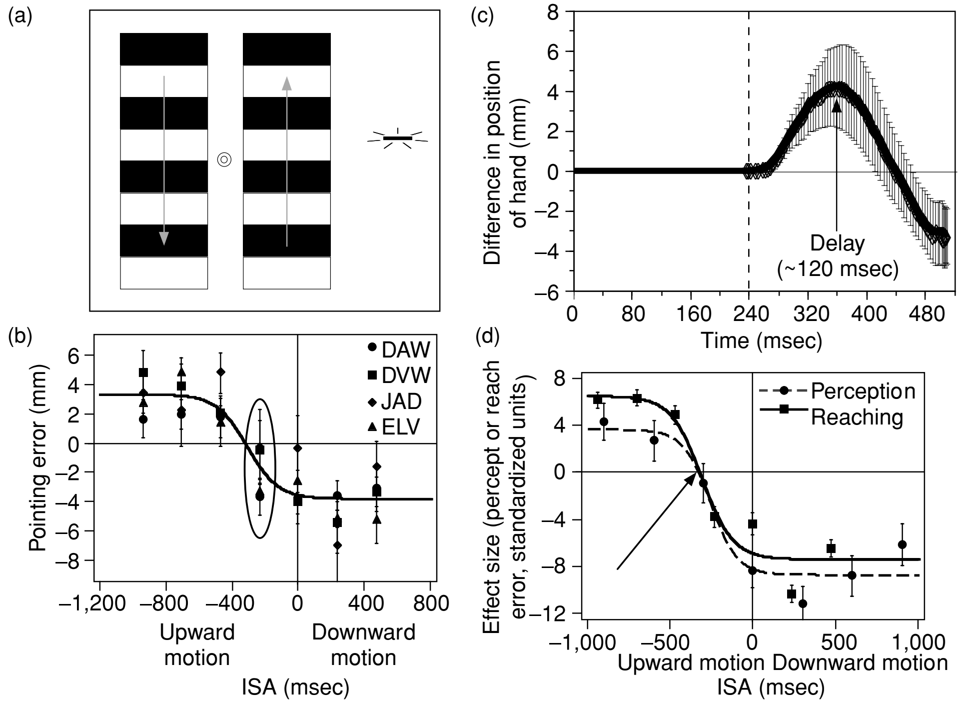


Fig. 9.3 Stimulus and results of an experiment demonstrating the influence of visual motion on fast visually guided reaching. (a) A static target was briefly flashed (rectangle) near a drifting grating. The nearby grating initially moved in one direction and then abruptly reversed direction. Subjects reached and hit the target with their index finger. (b) Endpoints of reaching movements. The abscissa shows the interstimulus asynchrony (ISA) between the target presentation and the motion reversal. Data are presented in a manner such that initially the grating drifted upward and then reversed direction and drifted downward. Targets presented well before (after) the motion reversal produced systematic upward (downward) errors in the endpoint of the reach, always in a direction consistent with that of the nearest visual motion. (c) For each given ISA (from (b)), the difference in the trajectory of the hand (vertical position over time) was calculated as a function of the direction of visual motion. The resulting difference is the modulation in the hand's position as a function of motion direction. At an ISA of ~−240 msec (c), the net effect of visual motion shows that the hand initially deviates upward, reaches a maximum deviation, and then deviates back downward, mimicking the motion reversal. The difference between the physical motion reversal and the moment of the hand's trajectory – the visuomotor delay – was about 120 msec (arrow). (d) The temporal tuning of visual motion's influence on perceived position (dashed line, circles) and its influence on reaching (solid line, squares). The perceived position of the flashed target (measured using Fig. 9.2(d)) depends on when it is presented relative to the grating's motion reversal (dashed line). Both the perceived position of the target and the endpoint position of the hand follow a similar time course. Figure adapted from Whitney et al. (2003), with permission.

the trajectory of the hand during both the programming and execution (online) phases of the reach. Consistent with the double-step studies described earlier, the visuomotor delay (the reaction time to a change in visual motion direction) was about 120–150 msec (arrow in Fig. 9.3(c)). This delay is quite brief, even though the target and influencing motion was separated by many degrees.

To test whether the deviation in the trajectory of the reaching movement depends on perceptual awareness of target position, Whitney et al. (2003) measured the perceived shift in the position of the target (as in Fig. 9.2(c)) using the same stimulus as in Fig. 9.3(d). When the time courses of the perceptual and motor effects are directly compared (Fig. 9.3(d)), it is clear that the deviation in the hand occurs as early as the deviation in the perceived position of the target. More importantly, the moment at which a perceptual decision is reached lags behind visual motion's influence on the reach. Figure 9.3(d) shows that the perceived position of the target depends on visual motion presented at least 300 msec after the target (targets presented at –300 msec ISA do not appear shifted upward despite the fact that they are presented during upward motion because their perceived position depends partially on downward visual motion presented more than 300 msec later). Reaching, however, has a delay of about 120 msec (Fig. 9.3(c)). Therefore, the visuomotor system does not depend upon a later perceptual decision but modifies the hand's trajectory online, whereas the perceived position of the target is still indeterminate. These results demonstrate that hand position cannot be based on the percept (because the direction of the hand's response precedes and is even opposite to the perceived position of the target). Of course, it is entirely possible that the same information is used by visuomotor and perceptual systems; however, it unfolds on different time scales. Regardless, *perceived* target position does not govern the direction or endpoint position of the hand. Other, nonperceptual representations of target position might be responsible for fast, online control of reaching.

9.3.2.3 Similarities between manual responses to object and background motion

The double-step studies described in Section 9.2 share many common traits with the influence of visual motion on reaching. First, both of the effects involve visual feedback used for online control, occur at very high speeds, and have comparable reaction times of around 100–150 msec (Brenner & Smeets 1997; Whitney et al. 2003; Saijo et al. 2005). In comparison, grip aperture reactions to changes in the visual shape or structure of objects are slower, as are other reactions to higher-level visual features (Paulignan et al. 1991a,b; Bock & Jungling 1999; Day & Lyon, 2000; Pisella et al. 2000).

As mentioned before, however, it is debated whether the visuomotor correction in the double-step paradigm is due to changes in target position per se or just the visual motion (or percept of motion) induced by the jumping target. One piece of evidence supporting the former (position matters, not just motion) comes from double-step studies in which the target was defined solely by chromatic differences (Pisella et al. 2000; Brenner & Smeets 2004). Color provides relatively weak input to the magnocellular stream (Livingstone & Hubel 1988; Derrington 2000), and color-defined motion is perceived more slowly

(Cropper & Derrington 1994, 1996), less accurately (e.g., motion standstill; Cavanagh et al. 1984; Cavanagh 1992; Lu & Sperling 1995; Lu et al. 1999), and, at least at low contrasts, with a position-based mechanism (Cropper & Derrington 1994; Seiffert & Cavanagh 1998). Therefore, if reactions to color changes are fast, this would be strong evidence that nonmoving features of objects (i.e., their positions alone) are sufficient to drive fast motor corrections. In one study using color-defined targets, Pisella et al. (2000) found that subjects were slow at overriding (halting) a reaching movement based on changes in the color of the target. This could indicate that visual motion matters for fast manual corrections. However, Brenner and Smeets (2004), measuring correcting responses rather than overriding responses, found that subjects were quite fast at correcting the direction of their pointing movement in response to changes in the color of a target. Therefore, visual motion may not be the only attribute of scenes and objects that can give rise to fast corrections; it may be that target position matters in both the double-step and background motion studies. Unfortunately, the issue is still unresolved because no study has successfully controlled for the luminance of the targets. The stimulus used by Brenner & Smeets (2004) was physically isoluminant, but was not equated psychophysically, so nonlinearities (Anstis & Cavanagh 1983; Lu & Sperling 2001b) could result in apparent brightness differences that could give rise to strong motion signals. Therefore, the double-step experiments need to be repeated with psychophysically equiluminant or second-order, contrast-defined stimuli that are detected by a position-based mechanism (Seiffert & Cavanagh 1998). These stimuli would provide definitive data on whether it is the position and/or velocity of targets in the double-step paradigm that drives the hand.

Another similarity between the influence of visual motion on reaching and the double-step paradigm is that both have revealed an important contribution of visual information during the programming as well as the execution (online) phases of the movement. A static target displaced prior to or during a reach influences the direction of the hand's trajectory (Section 9.2). Likewise, visual motion presented before or after a reaching movement is executed (during the programming or the execution phases) influences the direction of the hand (Whitney et al. 2003; Saijo et al. 2005; Gomi et al. 2006). The fact that visual motion during the programming phase of the movement influences a reach suggests that either the predicted reach position or the representation of the target may be shifted by visual motion just as it is when the target position is displaced in double-step experiments.

Although incomplete and necessitating further studies (see Section 9.5), the results reviewed here suggest a possible link between visual motion and double-step–induced modifications to the trajectory of the hand. If the link holds, this would suggest that the visuomotor system monitors target positions continuously (or at least frequently) and that visual motion causes updating of target positions similar to actual changes in their locations.

9.3.3 Following responses

Although it is suggested that the visual motion and double-step results are due to the manner in which the visuomotor system represents and updates target positions, there is another possibility. Visual motion (of the background or the target itself) could influence the hand directly. The "indirect" model assumes that a deviation in the trajectory of reaching movements in the presence of background motion occurs because the visuomotor system has an explicit representation of target position and updates, modifies, or shifts this position representation in the presence of background (or target) motion. Visual motion is important in this account, but only indirectly via an influence on the target positions tracked by the visuomotor system. An alternative is that the visuomotor system has direct access to visual motion, which influences a reach independent of the target representation. The most likely form of this alternative account is a manual following response (Saijo et al. 2005). The following discussion will draw a distinction between the direct (following response) and indirect (shifted target representation) hypotheses.

9.3.3.1 Background motion modulates smooth pursuit

Before examining the manual following response, it is worth considering the similar but distinct effects of background motion on the eye. It has long been known that the gain of pursuit eye movements (ratio of eye velocity to target velocity) is modulated by the presence and direction of background motion (Yee et al. 1983; Collewijn & Tamminga 1984; Keller & Khan 1986; Howard & Marton 1992; Masson et al. 1995; Mohrmann & Thier 1995; Zivotofsky et al. 1995; Niemann & Hoffmann 1997; Schwarz & Ilg 1999; Lindner et al. 2001, but c.f. Kowler et al. 1984). For example, pursuit gain can be reduced by upward of 10% when there is retinal motion opposite the direction of ocular pursuit (Masson et al. 1995), and this effect is specific to the energy or first-order motion in the display (Masson et al. 2002). Interestingly, however, gain modulation is a function of the distance between retinal motion and the pursuit target (Howard & Marton 1992; Masson et al. 2001; Goltz & Whitney 2004; Whitney & Goodale 2005); increasing the separation between the target and the motion by approximately 2 deg is sufficient to abolish the pursuit modulation (Goltz & Whitney 2004; Whitney & Goodale 2005). The influence of visual motion on reaching, however, operates over much larger distances, indicating differential spatial integration (or different gain or weighting functions for eye and reaching movements (Whitney et al. 2003; Saijo et al. 2005).

9.3.3.2 Ocular following response

In addition to eye movement speed (i.e., pursuit gain), it has also been shown that eye *position* is influenced by background retinal motion (Kawano & Miles 1986; Miles et al. 1986; Kawano et al. 1994; Masson et al. 2001), even in the absence of pursuit eye movements. This has been termed the "ocular following response." The latency of this response is much shorter than smooth pursuit eye movements because of a direct neural pathway between cortical regions of visual motion processing and subcortical areas responsible for the generation of eye movement commands (Shidara & Kawano 1993; Kawano et al. 1994).

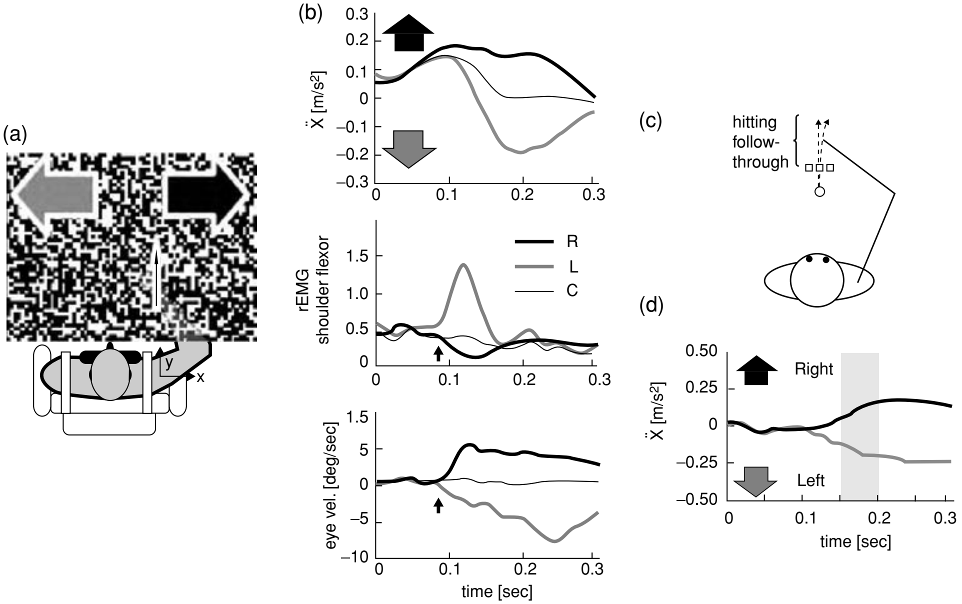


Fig. 9.4 Experimental setup to measure and quantify the manual following response, as well as arm kinetics, muscle, and ocular responses induced by visual motion. (a) Large field visual motion (right or left indicated by large arrows) was presented after the hand traveled 5 cm from its initial position. (b) Top panel: The hand's acceleration in the x-direction (orthogonal direction to the hand path) shows a deviation in the acceleration in the direction of the background motion within about 100 msec (the thick black, thick gray, and thin curves represent the rightward, leftward, and no visual motion conditions, respectively). Middle panel: The normalized and rectified electromyogram (EMG) of the shoulder flexor muscle. Bottom panel: Right eye velocity. The small vertical arrows in the middle and bottom panels show that the shoulder muscle response to visual motion is as fast as the ocular response. (c) Manual following response induced by a large-field visual motion during a follow-through movement, after hitting a target. The stop position of the reaching movement was not specified; subjects were instructed to follow through their reach in the most natural direction. (d) The averaged hand acceleration in the x-direction (the direction of the background visual motion) for those trials in which the background moved rightward or leftward during the follow-through phase of the reach. Figure adapted from Saijo et al. (2005), with permission.

9.3.3.3 Manual following response

The influence of visual motion on reaching movements (Brenner & Smeets 1997; Whitney et al. 2003; Saijo et al. 2005) is similar, in many respects, to the ocular following response (Gomi et al. 2005, 2006) and may therefore be characterized as a manual following response (Fig. 9.4). Akin to the ocular following response, the manual following response is a short-latency passive deviation in the trajectory of the hand in the direction of large-field visual motion (Brenner et al. 1997; Whitney et al. 2003; Saijo et al. 2005). This response can be elicited without a visible target, suggesting that positional information about the target itself may not be essential to response generation (Saijo et al. 2005).

The manual following response is distinct from several other reported effects. When tracking a moving target with the hand (manual pursuit), background texture (moving or stationary) influences the gain of the manual pursuit – but in a direction opposite that of the manual following response (Masson et al. 1995; Soechting et al. 2001). The manual following response is also not due to the location or movement of the eye (Sections 9.3.3.1 and 9.3.3.2). Although there are coordinated mechanisms serving both ocular and manual responses (Prablanc et al. 1979; Herman & Maulucci 1981; Biguer et al. 1982; Fischer & Rogal 1986; Turrell et al. 1998; Henriques et al. 1998b; Engel et al. 2000; Soechting et al. 2001; Ariff et al. 2002; Henriques et al. 2003), the manual following response induced by visual motion is opposite the pointing error caused by gaze shifts (Henriques et al. 1998b, 2003). More importantly, the latency of the arm muscle activity is comparable to that of the ocular following response (Saijo et al. 2005) as shown in Fig. 9.4(b). If the hand simply followed the eye, one would expect a more substantial delay. Finally, the ocular following response is eliminated when subjects fixate or when visual motion is presented in just one visual field; however, the manual following response remains strong in both of these cases (Whitney et al. 2003; Saijo et al. 2005). These observations diminish the possibility that eye–hand coordination produces the manual following response (see also Section 9.2).

Another possible explanation for motion's influence on reaching, as discussed previously, is that visual motion affects the representation of the target's position, which could lead to an online correction of the manual movement. Interestingly, however, the quick manual response can be induced without simultaneous presentation of the target and background visual motion (Saijo et al. 2005). More strikingly, the manual following response is still present during "follow-through" movements – even after the target is hit, the arm continues to deviate in the direction of background visual motion (Figs. 9.4(c) and (d)). Because the variation in the trajectory of follow-through movements after contact with the target is different from target-directed reaching, it has been suggested that the trajectory of the follow-through movement is distinct from the target-directed reaching phase (Saijo et al. 2005). From these observations, it appears that at least part of the quick manual response induced by a large-field visual motion is due to the direct influence of visual motion on the arm motor system rather than an indirect effect of visual motion on target representations.

The rapid manual following response occurs for visual motion not only in the fronto-vertical plane (Brenner & Smeets 1997; Whitney et al. 2003) but also in any direction along the horizontal plane (Saijo et al. 2004). This suggests that the manual following response is coded in 3D space rather than retinal 2D coordinates. Moreover, varying motion coherence (Newsome & Pare 1988) causes comparable changes in the following response (Saijo et al. 2005), implicating motion energy computations via motion processing units at a relatively late stage in the visual system – at least beyond the level of primary visual cortex, and most likely in visual area MT+. Using TMS, a recent study directly tested this hypothesis and found that stimulation of MT+ significantly reduced the manual following response (Whitney et al. 2007).

The observations reviewed here suggest that visual motion not only affects the representation of target position but also directly influences manual control. For the sake of efficiency, the visuomotor system analyzes information about the target and background in parallel. Simply disregarding background information would be difficult if not impossible, and for this reason it is necessary to consider how the visual system codes and uses background motion information. To elucidate the potentially diverse functional mechanisms that allow for skilled motor performance, it is important to compare and combine observations from both behavioral and physiological techniques. Together with the results above, a coherent story is beginning to emerge that visual motion information – even that of the background – is tightly linked to motor processing. This linkage operates through a direct influence of motion on reaching (manual following response) and may also involve a secondary indirect influence of visual motion on the represented positions of targets.

9.4 Beneficial uses of visual motion for reaching

Whether or not a distinction is drawn between (direct) motion-induced and (indirect) position-induced visuomotor responses, it is possible that both models operate under normal conditions and make comparable predictions in most circumstances. In fact, the influence of background motion on reaching could be an adaptive and beneficial response that the visuomotor system employs to guide reaching movements. Based on the statistical regularity of motion – the fact that background retinal motion is generally produced by self-movement (especially eye movements), the visuomotor system could access retinal motion as a means of gauging how the eye or body has moved relative to target objects (Whitney et al. 2003).

Although potentially adaptive, this explanation for the findings shown previously is counterintuitive. To reach to a target, the hand and the target must be coded in a common coordinate frame. If reach plans were coded in head, body, or hand-centered coordinates, then retinal motion's influence on reaching would be detrimental because the relative position of the hand and target do not change when the eyes move. A deviation in the hand's trajectory would therefore cause inaccuracy. However, there is evidence suggesting that targets are actually coded in eye-centered coordinates (Henriques et al. 1998a; Buneo et al. 2002; Crawford et al. 2004) and that coordinate transformations only happen on demand (Henriques et al. 1998a). If the common coordinate frame in which reach plans are coded is eye-centered, then every time we reach for an object, the visuomotor system must take into account eye position. Retinal motion in this scenario would be informative about changes in eye position and could be used to help update target (or eye) position for reaching.

9.4.1 Errors in reaching

If visual motion's influence on reaching is beneficial, then what is the error being corrected or compensated? One possibility is an error in extraretinal signals such as an efference copy. If an efference copy were perfect, the visuomotor system could update target positions continuously relative to gaze direction. Thus, the visuomotor system would have perfect information about relative eye-target positions and would accurately reach to objects. Unfortunately, efference copy signals are not always available, and even when they are available, they are routinely underestimated (i.e., the gain of the mechanism is less than one [Bridgeman 1995]). This is supported by demonstrations such as the Filehne and Aubert-Fleischl illusions (Mack & Herman 1973; Wertheim 1981; Freeman 2001). Underestimating eye movement amplitude (distance traveled) or velocity could cause reaching movements to systematically miss targets. This is especially true when targets are no longer visible after reach initiation. Visual motion could cause a bias in the trajectory of the reach opposite the direction of the eye movement, thereby reducing the systematic underestimation of efference copy. In fact, retinal motion opposite the direction of a pursuit eye movement improves reaching to the remembered position of a static target (Whitney & Goodale 2005).

Another possibility is that there are delays in coordinate transformations that necessitate a compensation mechanism. The target and hand must be coded in a common coordinate frame. Regardless of which frame this is, several transformations are required. Either the target is sequentially transformed from retinal to eye-, head-, body-, and hand-centered coordinates, or the hand is transformed in the reverse order. In both cases, unless transformations are instantaneous (a physical impossibility), there will be a mismatch between the target's physical location and the representation of the hand and target in a common coordinate frame. When reaching to a physically static object during a pursuit eye movement, the delays transforming the representations of the hand and target mean that by the time the reach is executed the eye will have continued along its trajectory, leaving a gap between the actual target location and the reach goal (i.e., an error in programming). Similar to the underestimation of an efference copy, this error would be reduced by retinal motion in a direction opposite that of the eye movement. When retinal motion is not available, on the other hand, this error does manifest itself (Whitney & Goodale 2005).

If the reach errors above are compensated by visual motion, then akinetopsic (motion blind; Zihl et al. 1983) patients might lack this compensation. In fact, these individuals are impaired when reaching to moving targets under free viewing conditions in which they move their eyes (Schenk et al. 2000), indicating that visual motion, and motion area MT+ in particular, is important for visually guided behavior (Schenk et al. 2005). Further studies of motion-blind patients, in which background and target motions are independently manipulated, would help establish the relevance of the background motion, per se. Given the speed with which visual motion is processed (Schmolesky et al. 1998) and that retinal motion correlates perfectly with eye movements, it would be enormously advantageous for the visuomotor system to use background visual motion as a source of feedback information for the control of visually guided reaching.

9.5 Future directions

A great deal of evidence suggests that low-level retinal motion (i.e., first-order luminance-defined motion) influences visually guided reaching. However, do other types of visual motion also exert an influence on reaching? For example, second-order motion (Cavanagh

& Mather 1989; Derrington et al. 2004) or other high-level motion such as transformational apparent motion, inferred motion, attentive tracking, or third-order motion (Cavanagh 1992; Hikosaka et al. 1993; Assad & Maunsell 1995; Watamaniuk & McKee 1995; Lu & Sperling 2001a; Tse & Logothetis 2002; Watamaniuk 2005) may or may not be used by the visuomotor system. Because these types of motion are thought to be processed in dorsal visual motion areas (Culham et al. 2001), if the visuomotor system receives direct and necessary input from these motion areas, we might expect reaching to depend on the percept of motion and not simply on the physical presence of retinal motion.

Further research is needed to determine whether the manual corrections in double-step studies are due, at least in part, to the visual motion produced by the jumping target. Although studies have begun to address this (Pisella et al. 2000; Brenner & Smeets 2004), the question remains. One way forward is to present a target object that moves in one direction while physically shifting in the opposite direction, pitting physical movement (motion energy) against position displacement, similar to what has been done in studies of pursuit (Masson et al. 2002). Ideally, however, the conflicting motion and position information would be specific to the target and not the background. Likewise, the conflicting information should be real – not illusory – as has been used in the past (Brenner & Smeets 1997). Although difficult to imagine, this is precisely what reverse-ϕ (Anstis 1970; Shioiri & Cavanagh 1990), four-stroke apparent motion (Anstis 1980) and the related missing fundamental illusion (Adelson & Bergen 1985) demonstrate.

9.6 Conclusions

Over the last several decades, the role of visual motion in visually guided action and motor control has been a topic of interest. Numerous studies have examined how eye movements and reaching are guided to moving objects; another group of independent but related studies has examined how posture and locomotion are guided by optic flow. More recently, a third group of studies has investigated the influence of background retinal motion on goal-directed reaching, manual tracking, and eye movements. Although distinct specific hypotheses are addressed by each of these studies, there is a broad consensus that visual motion – of objects themselves, scenes as a whole, or backgrounds alone – is sufficient to drive visuomotor behavior in a systematic, automatic, and fast manner. It is possible that all of these effects are determined by common neural mechanisms of motion processing and that these mechanisms serve one or more adaptive, beneficial roles. Future studies are therefore required, not only to establish the neural underpinnings of motion's influence on action but also to address the functional role(s) that retinal motion plays in motor control.

References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *J Opt Soc Am [A]* **2**(2): 284–299.

Anstis, S. M. (1970). Phi movement as a subtraction process. *Vision Res* **10**(12): 1411–1430.

Anstis, S. M. (1980). The perception of apparent movement. *Philos Trans R Soc Lond B Biol Sci* **290**(1038): 153–168.

Anstis, S. M., & Cavanagh, P. (1983). A minimum motion technique for judging equiluminance. In J. Mollon & R. T. Sharpe (eds.), *Color Vision: Physiology and Psychophysics* (155–166). London: Academic Press.

Ariff, G., Donchin, O., Nanayakkara, T., & Shadmehr, R. (2002). A real-time state predictor in motor control: study of saccadic eye movements during unseen reaching movements. *J Neurosci* **22**(17): 7721–7729.

Ashida, H. (2004). Action-specific extrapolation of target motion in human visual system. *Neuropsychologia* **42**(11): 1515–1524.

Assad, J. A., & Maunsell, J. H. (1995). Neuronal correlates of inferred motion in primate posterior parietal cortex. *Nature* **373**(6514): 518–521.

Ballard, D. H., Hayhoe, M. M., Li, F., & Whitehead, S. D. (1992). Hand-eye coordination during sequential tasks. *Philos Trans R Soc Lond B Biol Sci* **337**(1281): 331–338; discussion 338–339.

Biguer, B., Jeannerod, M., & Prablanc, C. (1982). The coordination of eye, head, and arm movements during reaching at a single visual target. *Exp Brain Res* **46**(2): 301–304.

Binsted, G., Chua, R., Helsen, W., & Elliott, D. (2001). Eye-hand coordination in goal-directed aiming. *Hum Mov Sci* **20**(4–5): 563–585.

Bock, O., & Jungling, S. (1999). Reprogramming of grip aperture in a double-step virtual grasping paradigm. *Exp Brain Res* **125**(1): 61–66.

Brenner, E., & Smeets, J. B. (1994). Different frames of reference for position and motion. *Naturwissenschaften* **81**(1): 30–32.

Brenner, E., & Smeets, J. B. (1997). Fast responses of the human hand to changes in target position. *J Mot Behav* **29**(4): 297–310.

Brenner, E., & Smeets, J. B. (2003). Fast corrections of movements with a computer mouse. *Spat Vis* **16**(3–4): 365–376.

Brenner, E., & Smeets, J. B. (2004). Colour vision can contribute to fast corrections of arm movements. *Exp Brain Res* **158**(3): 302–307.

Brenner, E., Smeets, J. B., & de Lussanet, M. H. (1998). Hitting moving targets. Continuous control of the acceleration of the hand on the basis of the target's velocity. *Exp Brain Res* **122**(4): 467–474.

Bridgeman, B. (1995). A review of the role of efference copy in sensory and oculomotor control systems. *Ann Biomed Eng* **23**(4): 409–422.

Bridgeman, B., Kirch, M., & Sperling, A. (1981). Segregation of cognitive and motor aspects of visual function using induced motion. *Percept Psychophys* **29**(4): 336–342.

Bridgeman, B., Lewis, S., Heit, G., & Nagle, M. (1979). Relation between cognitive and motor-oriented systems of visual position perception. *J Exp Psychol Hum Percept Perform* **5**(4): 692–700.

Bridgeman, B., Peery, S., & Anand, S. (1997). Interaction of cognitive and sensorimotor maps of visual space. *Percept Psychophys* **59**(3): 456–469.

Britten, K. H., & van Wezel, R. J. (1998). Electrical microstimulation of cortical area MST biases heading perception in monkeys. *Nat Neurosci* **1**(1): 59–63.

Buneo, C. A., Jarvis, M. R., Batista, A. P., & Andersen, R. A. (2002). Direct visuomotor transformations for reaching. *Nature* **416**(6881): 632–636.

Burr, D. C., & Ross, J. (2002). Direct evidence that “speedlines” influence motion mechanisms. *J Neurosci* **22**(19): 8661–8664.

Burr, D. C., Ross, J., & Morrone, M. C. (1986). Seeing objects in motion. *Proc R Soc Lond B Biol Sci* **227**(1247): 249–265.

Castiello, U., Paulignan, Y., & Jeannerod, M. (1991). Temporal dissociation of motor responses and subjective awareness. A study in normal subjects. *Brain* **114**(Pt 6): 2639–2655.

Cavanagh, P. (1992). Attention-based motion perception. *Science* **257**(5076): 1563–1565.

Cavanagh, P., & Mather, G. (1989). Motion: the long and short of it. *Spat Vis* **4**(2–3): 103–129.

Cavanagh, P., Tyler, C. W., & Favreau, O. E. (1984). Perceived velocity of moving chromatic gratings. *J Opt Soc Am A* **1**: 893–899.

Collewijn, H., & Tamminga, E. P. (1984). Human smooth and saccadic eye movements during voluntary pursuit of different target motions on different backgrounds. *J Physiol* **351**: 217–250.

Cowey, A., & Stoerig, P. (1991). The neurobiology of blindsight. *Trends Neurosci* **14**(4): 140–145.

Crawford, J. D., Medendorp, W. P., & Marotta, J. J. (2004). Spatial transformations for eye-hand coordination. *J Neurophysiol* **92**(1): 10–19.

Cropper, S. J., & Derrington, A. M. (1994). Motion of chromatic stimuli: first-order or second-order? *Vision Res* **34**(1): 49–58.

Cropper, S. J., & Derrington, A. M. (1996). Rapid colour-specific detection of motion in human vision. *Nature* **379**(6560): 72–74.

Culham, J., He, S., Dukelow, S., & Verstraten, F. A. (2001). Visual motion and the human brain: what has neuroimaging told us? *Acta Psychol (Amst)* **107**(1–3): 69–94.

Dassonville, P., Bridgeman, B., Kaur Bala, J., Thiem, P., & Sampanes, A. (2004). The induced Roelofs effect: two visual systems or the shift of a single reference frame? *Vision Res* **44**(6): 603–611.

Day, B. L., & Lyon, I. N. (2000). Voluntary modification of automatic arm movements evoked by motion of a visual target. *Exp Brain Res* **130**(2): 159–168.

De Valois, R. L., & De Valois, K. K. (1991). Vernier acuity with stationary moving gabors. *Vision Res* **31**(9): 1619–1626.

Del Viva, M. M., & Morrone, M. C. (1998). Motion analysis by feature tracking. *Vision Res* **38**(22): 3633–3653.

Derrington, A. M. (2000). Vision: can colour contribute to motion? *Curr Biol* **10**(7): R268–270.

Derrington, A. M., Allen, H. A., & Delicato, L. S. (2004). Visual mechanisms of motion analysis and motion perception. *Annu Rev Psychol* **55**: 181–205.

Desmurget, M., & Grafton, S. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends Cogn Sci* **4**(11): 423–431.

Diedrichsen, J., Nambisan, R., Kennerley, S. W., & Ivry, R. B. (2004). Independent on-line control of the two hands during bimanual reaching. *Eur J Neurosci* **19**(6): 1643–1652.

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. *J Neurophysiol* **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. *J Neurophysiol* **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. *J Neurophysiol* **86**(4): 1991–2000.

Engel, K. C., Anderson, J. H., & Soechting, J. F. (2000). Similarity in the response of smooth pursuit and manual tracking to a change in the direction of target motion. *J Neurophysiol* **84**(3): 1149–1156.

Fischer, B., & Rogal, L. (1986). Eye-hand-coordination in man: a reaction time study. *Biol Cybern* **55**(4): 253–261.

Freeman, T. C. (2001). Transducer models of head-centred motion perception. *Vision Res* **41**(21): 2741–2755.

Geisler, W. S. (1999). Motion streaks provide a spatial code for motion direction. *Nature* **400**(6739): 65–69.

Gibson, J. J. (1986). *The Ecological Approach to Visual Perception*. Hillsdale, NJ: Erlbaum.

Goltz, H. C., & Whitney, D. (2004). The influence of background motion on smooth pursuit: separation matters. *Journal of Vision* **4**: 649.

Gomi, H., Abekawa, N., & Nishida, S. (2005). Implicit sensorimotor control: rapid motor responses of arm and eye share the visual motion encoding [Abstract]. *Journal of Vision* **5**(8): 363a; http://journalofvision.org/5/8/363/, doi:10.1167/5.8.363

Gomi, H., Abekawa, N., & Nishida, S. (2006). Spatiotemporal tuning of rapid interactions between visual-motion analysis and reaching movement. *J Neurosci* **26**(20): 5301–5308.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends Neurosci* **15**(1): 20–25.

Goodale, M. A., Pelisson, D., & Prablanc, C. (1986). Large adjustments in visually guided reaching do not depend on vision of the hand or perception of target displacement. *Nature* **320**(6064): 748–750.

Gray, R. (2002). Behavior of college baseball players in a virtual batting task. *J Exp Psychol Hum Percept Perform* **28**(5): 1131–1148.

Greenlee, M. W. (2000). Human cortical areas underlying the perception of optic flow: brain imaging studies. *Int Rev Neurobiol* **44**: 269–292.

Hayes, A. (2000). Apparent position governs contour-element binding by the visual system. *Proc R Soc Lond B Biol Sci* **267**(1450): 1341–1345.

Henriques, D. Y., Klier, E. M., Smith, M. A., Lowy, D., & Crawford, J. D. (1998a). Gaze-centered remapping of remembered visual space in an open-loop pointing task. *J Neurosci* **18**(4): 1583–1594.

Henriques, D. Y., Klier, E. M., Smith, M. A., Lowy, D., & Crawford, J. D. (1998b). Gaze-centered remapping of remembered visual space in an open-loop pointing task. *J Neurosci* **18**(4): 1583–1594.

Henriques, D. Y., Medendorp, W. P., Gielen, C. C., & Crawford, J. D. (2003). Geometric computations underlying eye-hand coordination: orientations of the two eyes and the head. *Exp Brain Res* **152**(1): 70–78.

Herman, R., & Maulucci, R. (1981). Visually triggered eye-arm movements in man. *Exp Brain Res* **42**(3–4): 392–398.

Hikosaka, K., Iwai, E., Saito, H., & Tanaka, K. (1988). Polysensory properties of neurons in the anterior bank of the caudal superior temporal sulcus of the macaque monkey. *J Neurophysiol* **60**(5): 1615–1637.

Hikosaka, O., Miyauchi, S., & Shimojo, S. (1993). Focal visual attention produces illusory temporal order and motion sensation. *Vision Res* **33**(9): 1219–1240.

Howard, I. P., & Marton, C. (1992). Visual pursuit over textured backgrounds in different depth planes. *Exp Brain Res* **90**(3): 625–629.

Huk, A. C., Dougherty, R. F., & Heeger, D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. *J Neurosci* **22**(16): 7195–7205.

Ingle, D. (1973). Two visual systems in the frog. *Science* **181**(104): 1053–1055.

Jacob, P., & Jeannerod, M. (2003). *Ways of Seeing*. Oxford: Oxford University Press.

Kawano, K., & Miles, F. A. (1986). Short-latency ocular following responses of monkey. II. Dependence on a prior saccadic eye movement. *J Neurophysiol* **56**(5): 1355–1380.

Kawano, K., Shidara, M., Watanabe, Y., & Yamane, S. (1994). Neural activity in cortical area MST of alert monkey during ocular following responses. *J Neurophysiol* **71**(6): 2305–2324.

Keller, E. L., & Khan, N. S. (1986). Smooth-pursuit initiation in the presence of a textured background in monkey. *Vision Res* **26**(6): 943–955.

Kerzel, D., & Gegenfurtner, K. R. (2003). Neuronal processing delays are compensated in the sensorimotor branch of the visual system. *Curr Biol* **13**(22): 1975–1978.

Kerzel, D., & Gegenfurtner, K. R. (2005). Motion-induced illusory displacement reexamined: differences between perception and action? *Exp Brain Res* **162**(2): 191–201.

Kowler, E., van der Steen, J., Tamminga, E. P., & Collewijn, H. (1984). Voluntary selection of the target for smooth eye movement in the presence of superimposed, full-field stationary and moving stimuli. *Vision Res* **24**(12): 1789–1798.

Land, M. F., & McLeod, P. (2000). From eye movements to actions: how batsmen hit the ball. *Nat Neurosci* **3**(12): 1340–1345.

Lee, D. N. (1980). The optic flow field: the foundation of vision. *Philos Trans R Soc Lond B Biol Sci* **290**(1038): 169–179.

Lee, D. N., & Aronson, E. (1974). Visual proprioceptive control of standing in human infants. *Perception & Psychophysics* **15**: 529–532.

Lee, D. N., & Reddish, P. E. (1981). Plummeting gannets: a paradigm of ecological optics. *Nature* **293**(5830): 293–294.

Lindner, A., Schwarz, U., & Ilg, U. J. (2001). Cancellation of self-induced retinal image motion during smooth pursuit eye movements. *Vision Res* **41**(13): 1685–1694.

Livingstone, M., & Hubel, D. (1988). Segregation of form, colour, movement, and depth: anatomy, physiology, and perception. *Science* **240**(4853): 740–749.

Lu, Z. L., Lesmes, L. A., & Sperling, G. (1999). Perceptual motion standstill in rapidly moving chromatic displays. *Proc Natl Acad Sci U S A* **96**(26): 15374–15379.

Lu, Z. L., & Sperling, G. (1995). Attention-generated apparent motion. *Nature* **377**(6546): 237–239.

Lu, Z. L., & Sperling, G. (2001a). Three-systems theory of human visual motion perception: review and update. *J Opt Soc Am A Opt Image Sci Vis* **18**(9): 2331–2370.

Lu, Z. L., & Sperling, G. (2001b). Sensitive calibration and measurement procedures based on the amplification principle in motion perception. *Vision Res* **41**(18): 2355–2374.

Mack, A., & Herman, E. (1973). Position constancy during pursuit eye movement: an investigation of the Filehne illusion. *Q J Exp Psychol* **25**(1): 71–84.

Masson, G., Proteau, L., & Mestre, D. R. (1995). Effects of stationary and moving textured backgrounds on the visuo- oculo-manual tracking in humans. *Vision Res* **35**(6): 837–852.

Masson, G. S., Busettini, C., Yang, D. S., & Miles, F. A. (2001). Short-latency ocular following in humans: sensitivity to binocular disparity. *Vision Res* **41**(25–26): 3371–3387.

Masson, G. S., Yang, D. S., & Miles, F. A. (2002). Reversed short-latency ocular following. *Vision Res* **42**(17): 2081–2087.

McGraw, P. V., Whitaker, D., Skillen, J., & Chung, S. T. (2002). Motion adaptation distorts perceived visual position. *Curr Biol* **12**(23): 2042–2047.

Miles, F. A., Kawano, K., & Optican, L. M. (1986). Short-latency ocular following responses of monkey. I. Dependence on temporospatial properties of visual input. *J Neurophysiol* **56**(5): 1321–1354.

Mohrmann, H., & Thier, P. (1995). The influence of structured visual backgrounds on smooth-pursuit initiation, steady-state pursuit and smooth-pursuit termination. *Biol Cybern* **73**(1): 83–93.

Mohrmann-Lendla, H., & Fleischer, A. G. (1991). The effect of a moving background on aimed hand movements. *Ergonomics* **34**(3): 353–364.

Nakayama, K., & Tyler, C. W. (1981). Psychophysical isolation of movement sensitivity by removal of familiar position cues. *Vision Res* **21**(4): 427–433.

Neggers, S. F., & Bekkering, H. (2001). Gaze anchoring to a pointing target is present during the entire pointing movement and is driven by a non-visual signal. *J Neurophysiol* **86**(2): 961–970.

Newsome, W. T., & Pare, E. B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *J Neurosci* **8**(6): 2201–2211.

Niemann, T., & Hoffmann, K. P. (1997). The influence of stationary and moving textured backgrounds on smooth-pursuit initiation and steady state pursuit in humans. *Exp Brain Res* **115**(3): 531–540.

Nishida, S. (2004). Motion-based analysis of spatial patterns by the human visual system. *Curr Biol* **14**(10): 830–839.

Nishida, S., & Johnston, A. (1999). Influence of motion signals on the perceived position of spatial pattern. *Nature* **397**(6720): 610–612.

Paillard, J. (1982). The contribution of peripheral and central vision to visually guided reaching. In D. J. Ingle, M. A. Goodale, & D. J. W. Mansfield (eds.), *Analysis of Visual Behaviour* (367–385). Cambridge, MA: MIT Press.

Paillard, J. (1996). Fast and slow feedback loops for the visual correction of spatial errors in a pointing task: A reappraisal. *Can J Physiol Pharmacol* **74**(4): 401–417.

Paulignan, Y., Jeannerod, M., MacKenzie, C., & Marteniuk, R. (1991a). Selective perturbation of visual input during prehension movements. 2. The effects of changing object size. *Exp Brain Res* **87**(2): 407–420.

Paulignan, Y., MacKenzie, C., Marteniuk, R., & Jeannerod, M. (1991b). Selective perturbation of visual input during prehension movements. 1. The effects of changing object position. *Exp Brain Res* **83**(3): 502–512.

Pelisson, D., Prablanc, C., Goodale, M. A., & Jeannerod, M. (1986). Visual control of reaching movements without vision of the limb. II. Evidence of fast unconscious processes correcting the trajectory of the hand to the final position of a double-step stimulus. *Exp Brain Res* **62**(2): 303–311.

Pelz, J., Hayhoe, M., & Loeber, R. (2001). The coordination of eye, head, and hand movements in a natural task. *Exp Brain Res* **139**(3): 266–277.

Pisella, L., Grea, H., Tilikete, C., Vighetto, A., Desmurget, M., Rode, G., et al. (2000). An 'automatic pilot' for the hand in human posterior parietal cortex: toward reinterpreting optic ataxia. *Nat Neurosci* **3**(7): 729–736.

Post, R. B., & Welch, R. B. (2004). Studies of open-loop pointing in the presence of induced motion. *Percept Psychophys* **66**(6): 1045–1055.

Prablanc, C., Echallier, J. F., Komilis, E., & Jeannerod, M. (1979). Optimal response of eye and hand motor systems in pointing at a visual target. I. Spatio-temporal characteristics of eye and hand movements and their relationships when varying the amount of visual information. *Biol Cybern* **35**(2): 113–124.

Prablanc, C., & Martin, O. (1992). Automatic control during hand reaching at undetected two-dimensional target displacements. *J Neurophysiol* **67**(2): 455–469.

Previc, F. H. (1992). The effects of dynamic visual stimulation on perception and motor control. *J Vestib Res* **2**(4): 285–295.

Proteau, L., & Masson, G. (1997). Visual perception modifies goal-directed movement control: supporting evidence from a visual perturbation paradigm. *Q J Exp Psychol A* **50**(4): 726–741.

Ramachandran, V. S., & Anstis, S. M. (1990). Illusory displacement of equiluminous kinetic edges. *Perception* **19**(5): 611–616.

Regan, D. (1997). Visual factors in hitting and catching. *J Sports Sci* **15**(6): 533–558.

Saijo, N., Murakami, I., Nishida, S., & Gomi, H. (2005). Large-field visual motion directly induces an involuntary rapid manual following response. *J Neurosci* **25**(20): 4941–4951.

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. *J Neurosci* **6**(1): 145–157.

Savelsbergh, G. J., Whiting, H. T., & Bootsma, R. J. (1991). Grasping tau. *J Exp Psychol Hum Percept Perform* **17**(2): 315–322.

Schenk, T., Ellison, A., Rice, N., & Milner, A. D. (2005). The role of v5/MT+ in the control of catching movements: an RTMS study. *Neuropsychologia* **43**(2): 189–198.

Schenk, T., Mai, N., Ditterich, J., & Zihl, J. (2000). Can a motion-blind patient reach for moving objects? *Eur J Neurosci* **12**(9): 3351–3360.

Schenk, T., Mair, B., & Zihl, J. (2004). The use of visual feedback and on-line target information in catching and grasping. *Exp Brain Res* **154**(1): 85–96.

Schmolesky, M. T., Wang, Y., Hanes, D. P., Thompson, K. G., Leutgeb, S., Schall, J. D., et al. (1998). Signal timing across the macaque visual system. *J Neurophysiol* **79**(6): 3272–3278.

Schneider, G. E. (1969). Two visual systems. *Science* **163**(870): 895–902.

Schwarz, U., & Ilg, U. J. (1999). Asymmetry in visual motion processing. *Neuroreport* **10**(12): 2477–2480.

Seiffert, A. E., & Cavanagh, P. (1998). Position displacement, not velocity, is the cue to motion detection of second-order stimuli. *Vision Res* **38**(22): 3569–3582.

Sheth, B. R., & Shimojo, S. (2000). In space, the past can be recast but not the present. *Perception* **29**(11): 1279–1290.

Shidara, M., & Kawano, K. (1993). Role of Purkinje cells in the ventral paraflocculus in short-latency ocular following responses. *Exp Brain Res* **93**: 185–195.

Shioiri, S., & Cavanagh, P. (1990). Isi produces reverse apparent motion. *Vision Res* **30**(5): 757–768.

Smeets, J. B., & Brenner, E. (1995a and b). Perception and action are based on the same visual information: distinction between position and velocity. *J Exp Psychol Hum Percept Perform* **21**(1): 19–31.

Soechting, J. F., Engel, K. C., & Flanders, M. (2001). The Duncker illusion and eye-hand coordination. *J Neurophysiol* **85**(2): 843–854.

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. *J Neurophysiol* **62**(3): 626–641.

Tresilian, J. R. (1993). Four questions of time to contact: a critical examination of research on interceptive timing. *Perception* **22**(6): 653–680.

Trevarthen, C. B. (1968). Two mechanisms of vision in primates. *Psychol Forsch* **31**(4): 299–348.

Tse, P. U., & Logothetis, N. K. (2002). The duration of 3-d form analysis in transformational apparent motion. *Percept Psychophys* **64**(2): 244–265.

Turrell, Y., Bard, C., Fleury, M., Teasdale, N., & Martin, O. (1998). Corrective loops involved in fast aiming movements: effect of task and environment. *Exp Brain Res* **120**(1): 41–51.

Ullman, S. (1979). *The Interpretation of Visual Motion*. Cambridge, MA: MIT Press.

van Asten, W. N., Gielen, C. C., & van der Gon, J. J. (1988). Postural movements induced by rotations of visual scenes. *J Opt Soc Am A* **5**(10): 1781–1789.

van Santen, J. P. H., & Sperling, G. (1985). Elaborated Reichard detectors. *J Opt Soc Am A* **2**(2): 300–321.

van Sonderen, J. F., Denier van der Gon, J. J., & Gielen, C. C. (1988). Conditions determining early modification of motor programmes in response to changes in target location. *Exp Brain Res* **71**(2): 320–328.

van Sonderen, J. F., Gielen, C. C., & Denier van der Gon, J. J. (1989). Motor programmes for goal-directed movements are continuously adjusted according to changes in target location. *Exp Brain Res* **78**(1): 139–146.

Wang, Y., & Frost, B. J. (1992). Time to collision is signalled by neurons in the nucleus rotundus of pigeons. *Nature* **356**(6366): 236–238.

Warren, W. H., Jr., Kay, B. A., Zosh, W. D., Duchon, A. P., & Sahuc, S. (2001). Optic flow is used to control human walking. *Nat Neurosci* **4**(2): 213–216.

Watamaniuk, S. N. (2005). The predictive power of trajectory motion. *Vision Res* **45**(24): 2993–3003.

Watamaniuk, S. N., & McKee, S. P. (1995). Seeing motion behind occluders. *Nature* **377**(6551): 729–730.

Watson, A. B., & Ahumada, A. J., Jr. (1985). Model of human visual-motion sensing. *J Opt Soc Am A* **2**(2): 322–341.

Wertheim, A. H. (1981). On the relativity of perceived motion. *Acta Psychol (Amst)* **48**(1–3): 97–110.

Whitaker, D., McGraw, P. V., & Pearson, S. (1999). Non-veridical size perception of expanding and contracting objects. *Vision Res* **39**(18): 2999–3009.

Whitney, D. (2002). The influence of visual motion on perceived position. *Trends Cogn Sci* **6**(5): 211–216.

Whitney, D., & Cavanagh, P. (2000). Motion distorts visual space: shifting the perceived position of remote stationary objects. *Nat Neurosci* **3**(9): 954–959.

Whitney, D., Ellison, A., Rice, N. J., Arnold, D., Goodale, M., Walsh, V., et al. (2007). Visually guided reaching depends on motion area MT+. *Cereb Cortex* **17**(11): 2644–2649.

Whitney, D., & Goodale, M. A. (2005). Visual motion due to eye movements helps guide the hand. *Exp Brain Res* **162**(3): 394–400. Epub 2005 Jan 2015.

Whitney, D., Westwood, D. A., & Goodale, M. A. (2003). The influence of visual motion on fast reaching movements to a stationary object. *Nature* **423**(6942): 869–873.

Yamagishi, N., Anderson, S. J., & Ashida, H. (2001). Evidence for dissociation between the perceptual and visuomotor systems in humans. *Proc R Soc Lond B Biol Sci* **268**(1470): 973–977.

Yee, R. D., Daniels, S. A., Jones, O. W., Baloh, R. W., & Honrubia, V. (1983). Effects of an optokinetic background on pursuit eye movements. *Invest Ophthalmol Vis Sci* **24**(8): 1115–1122.

Zihl, J., von Cramon, D., & Mai, N. (1983). Selective disturbance of movement vision after bilateral brain damage. *Brain* **106**(Pt 2): 313–340.

Zivotofsky, A. Z., Averbuch-Heller, L., Thomas, C. W., Das, V. E., Discenna, A. O., & Leigh, R. J. (1995). Tracking of illusory target motion: differences between gaze and head responses. *Vision Res* **35**(21): 3029–3035.

Part II

Temporal phenomena: perception

10

Saccadic chronostasis and the continuity of subjective temporal experience across eye movements

KIELAN YARROW, PATRICK HAGGARD, AND JOHN C. ROTHWELL

Summary

The term "saccadic chronostasis" refers to the subjective temporal lengthening of a visual stimulus perceived following a saccadic eye movement. In this chapter, we discuss our preferred account of the illusion, which posits that the onset of the postsaccadic stimulus is antedated to a moment just prior to movement initiation, and review supporting evidence that illustrates key characteristics of the illusion, including its dependency on saccade extent. We conclude with a brief discussion of other examples of biased time perception that have been linked to saccadic chronostasis.

10.1 Introduction

When people make a saccadic eye movement to fixate a new visual target, they overestimate the duration for which that target is perceived (Yarrow et al. 2001). This illusion, which we have called saccadic chronostasis, has been demonstrated using the following basic procedure. Subjects make a saccade to a target that changes form or color during the saccade. They judge the duration of the new target stimulus relative to subsequently presented reference stimuli, and these judgments are used to determine a point of subjective equality (PSE; the point at which the target and reference stimuli are perceived to have identical durations). This procedure is schematized in Fig. 10.1. The same task performed while fixating forms a control. Reduced PSEs in saccadic conditions compared to control fixation conditions are a gauge of the temporal overestimation of the postsaccadic stimulus.

A similar effect can also be observed in a more everyday setting. The "stopped clock" illusion occurs when upon glancing at a watch with a moving second hand we think, just for a moment, that it has stopped working. This experience is one that many people recognize and prompted the first investigations of saccadic chronostasis. It does not occur every time we look at our watch, but only on those occasions when the watch hand (or a digital counter) changes just before or during the saccade (Brown & Rothwell 1997). In these cases, the next interval seems to exceed the prescribed duration of one second.

Aside from explaining this common perceptual experience, why study saccadic chronostasis? In this chapter, we propose that the illusion helps explain how our visual experience

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

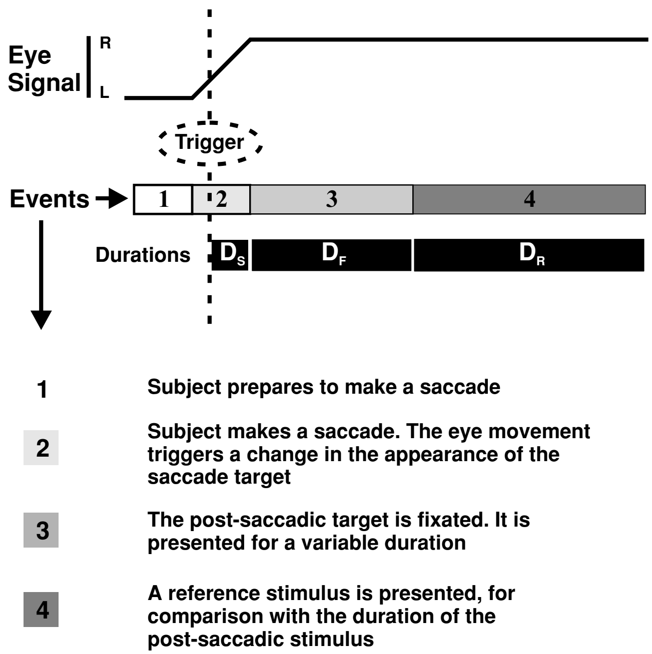


Fig. 10.1 Schematic of the saccade condition from a typical experimental task. D_S represents the duration for which the postsaccadic target stimulus is displayed on screen during the saccade. D_F represents the duration for which it is subsequently fixated. D_R represents the duration for which a reference stimulus is presented. The display time of the postsaccadic stimulus ($D_S + D_F$) is varied from trial to trial. A consistent finding is that subjects feel as though they have seen the postsaccadic stimulus and the reference stimulus for identical durations when D_F is significantly lower than D_R.

consists of a seamless progression of fixations without any intervening saccadic gaps. In elucidating this account, our approach will be as follows. First, we will describe a number of key findings from the various saccadic chronostasis experiments we have conducted to date, followed in each case by our interpretation of them. Next, we will discuss some methodological points that bear on the interpretation of saccadic chronostasis experiments. Finally, we will briefly discuss some other biases in temporal perception that have been linked to saccadic chronostasis and assess commonalities and distinguishing features.

10.2 Key experimental results supporting the antedating account

10.2.1 Saccadic chronostasis is greater for longer saccades than for shorter saccades

Using the basic methodology described previously, Yarrow et al. (2001) found that subjects did indeed overestimate the duration of a stimulus they had just fixated with a rapid eye movement. Subjects made saccades of either 22 deg or 55 deg extent and judged the duration of a postsaccadic stimulus. They made the same judgment in two control conditions involving fixation at an identical orbital eccentricity. The size of the resultant bias was found to depend upon the duration of the saccade. The bias was greater in the large saccade condition than in the small saccade condition, and this difference was comparable to the difference in saccadic duration. We will refer to this as the *saccade length* effect.

This finding is consistent with the hypothesis that the timeline of events recalled following a saccade is nonveridical. Observers do not report a duration that is consistent with having perceived the postsaccadic stimulus at the moment it was foveated (the end of the saccade) or even at the moment it first appeared (during the saccade). Instead, they report a duration that is consistent with having seen this stimulus approximately 50 msec *before* they moved their eyes. We refer to this as the *antedating* hypothesis. Note that antedating might result from either prospective or retrospective processes. It is possible that a clock process is automatically initiated before the onset of each saccade (a prospective account). On the other hand, events occurring around the time of the saccade could be retrospectively interpreted before a judgment about duration is made. In Sections 10.2.4 and 10.2.5 we will make a tentative case suggesting that timing processes are initiated prospectively but that the reported experience reflects a combination of this prospective time estimate with detailed visual information available only in retrospect.

There are a number of other observations that can inform, or be interpreted within, the antedating framework. During a saccade, visual input is highly degraded. High spatial frequency visual information is smeared by the rapid movement of the eye, whereas low spatial frequency visual information is subjected to an active process of saccadic suppression (Ross et al. 2001). The visual input is further suppressed as a result of backward masking by the postsaccadic image (Campbell & Wurtz 1978). These results explain our failure to perceive motion during a saccade but not our failure to experience any interruption of normal vision during this interval. The recollection of a timeline of events that effectively eliminates the saccade seems to provide the final piece in this puzzle, explaining the complete lack of visual experience during the period our eyes are in motion.

10.2.2 Stimulus duration does not influence the magnitude of saccadic chronostasis

In two experiments, Yarrow et al. (2004a) selected a range of reference durations between 100 msec and 1333 msec and then determined the PSE for a postsaccadic stimulus for each of them. They found that the size of the saccadic chronostasis effect was constant and independent of stimulus duration.

We were motivated to run these experiments because although the saccade length effect is consistent with the antedating hypothesis, it is not conclusive on its own. Problems of interpretation arise because the measure being used (the perceived duration of the postsaccadic stimulus) cannot be unambiguously related to the perceptual event about which we are making a claim (the onset of the postsaccadic target). In physics, the duration of an interval can only be changed by adjusting the onset time of events that border the interval. Psychologically, however, this is not the case. Perceived duration can be affected by a number of nontemporal factors (Allan 1979). Many theorists relate these changes in perceived time to the rate at which some hypothetical internal clock is functioning (Treisman et al. 1990; Wearden et al. 1998). Hence our earlier results could reflect a change in clock rate rather than temporal antedating (Hodinott-Hill et al. 2002).

If saccadic chronostasis is the result of a change in clock rate, the size of the effect should depend on the duration of the postsaccadic interval being judged. This follows because subjective time will equal objective time multiplied by clock rate. This prediction was not verified. However, this result does not conclusively rule out an account based upon a change in clock rate if we accept that such a change might be extremely transient, that is, already complete by the time the shortest duration stimulus that was tested had terminated. Hence the data obtained merely constrain a clock rate account, implying an increase that is both dramatic and brief.

10.2.3 The subjective time of onset for a postsaccadic stimulus is much earlier than for the same stimulus judged at fixation

Yarrow et al. (2006a) carried out a typical saccadic chronostasis experiment in which subjects made saccades of either 10 deg or 50 deg extent. The same subjects also completed an experiment in which a brief auditory stimulus (a beep) sounded around the time they moved their eyes. In this case, their task was to judge whether the auditory stimulus came on before or after they first saw the postsaccadic visual stimulus (i.e., a cross-modal temporal order judgment). A large bias emerged in both experiments. In the standard chronostasis experiment, the postsaccadic stimulus had an extended subjective duration compared to control conditions. In the temporal order judgment experiment, the beep had to be sounded before the postsaccadic target was foveated in order to be perceived as synchronous with it. Importantly, this bias was assessed relative to a control condition without an eye movement to take account of the latency differences between the visual and auditory modalities.

These experiments provide more direct evidence supporting the antedating account. In both experiments, effects were larger following large saccades than following small saccades. Hence, two completely different tasks applied to the same experimental situation provide consistent evidence that the perceived time of onset for the postsaccadic stimulus was earlier than both the moment it was foveated and its physical onset on the screen. The temporal order judgment task is explicitly an event judgment task. It therefore circumvents the problem of measuring the chronostasis effect previously using interval judgments as an implicit index of subjective events.

10.2.4 Saccadic chronostasis is eliminated when the saccade target jumps during a saccade

Yarrow et al. (2001) found that saccadic chronostasis did not occur when the saccade target was noticeably displaced (i.e., jumped horizontally by around 3 deg) at the same time it changed form during the saccade. The saccadic chronostasis effect returned partially when the same displacement went unnoticed and was observed as usual when distracters appeared close to the saccade target during the saccade. More recent unpublished work suggests that a second object in the postsaccadic display other than the saccade target can be subject to saccadic chronostasis when participants judge its duration. Just as for the saccade target object, displacement of this second object also eliminates saccadic chronostasis.

Antedating might be implemented in a prospective or retrospective fashion. In nonlaboratory settings, the stopped clock illusion is frequently experienced despite the fact that the postsaccadic target is both unpredictable and available only to peripheral vision prior to the saccade. This would seem to preclude a detailed real-time percept at the time of the target's perceived (presaccadic) onset and favor a retrospective account. However, the timing mechanism might still be prospective, with the subsequent detailed percept being retrospectively anchored to a moment determined in real time. Of course, anecdotal reports of the stopped clock illusion lack proper experimental control. In the lab, the postsaccadic target is highly predictable, so anticipation can contribute to perisaccadic perception, implying that both timing processes and conscious perception might be determined prospectively.

The fact that chronostasis disappears when the saccade target jumps seems to imply a retrospective interpretive process because perception changes depending upon an event that occurs only after the target is reported to be seen. High-level processes such as these are appropriate when sensory input is degraded or ambiguous and may depend upon prior expectations (e.g., Yang & Purves 2003). For perisaccadic vision, one reasonable expectation is that the world has not changed a great deal during the period of the saccade. When we designed our experiments with a jumping target, we predicted that perception would be modulated when sensory evidence is available that contradicts the expectation of perisaccadic continuity. The target's jump may violate expectations about the stability of the external world across eye movements and therefore veto a default tendency to antedate the postsaccadic stimulus.

Our experiments reveal the timeline of events that subjects recall across a saccade, but a question remains about exactly what they believe they have actually seen in the saccadic interval. We have recently begun to address this issue. We observed chronostasis following saccades to a moving object, but our subjects did not perceive a corresponding period of stimulus motion filling the saccadic gap (Yarrow et al. 2006b). Their percept was inferred based upon the first position at which they reported seeing the postsaccadic target, which was actually ahead of its true position, not behind it (although slightly less ahead of its true position than a similar target judged without a prior saccade). It appears then that our perception of the timing of events can be adjusted without requiring a complementary adjustment to spatial vision; an example of how different stimulus properties can become dissociated in conscious perception.

10.2.5 Saccadic chronostasis is obtained equally for different kinds of saccades

In two experiments, Yarrow et al. (2004b) found that the saccadic chronostasis effect could be obtained with a similar magnitude for many different kinds of saccades, including self-timed saccades, pro- and antisaccades, and even express saccades.

These experiments offer some insight into the possible neural locus of the effect. An extensive network of brain areas is involved in the production of saccades, but express saccades (those elicited in a gap paradigm with a latency of 70–130 msec; Fischer & Ramsperger 1984) are generally held to be generated in exclusively subcortical regions (Hopp & Fuchs 2002). The antedating hypothesis predicts that an efference copy signal relating to the saccade must be transmitted to brain regions that can determine when the postsaccadic stimulus appeared, and for how long it was presented. The finding of saccadic chronostasis following express saccades suggests that this signal may originate in the superior colliculus. We also found that chronostasis occurs with a similar magnitude for pro- and antisaccades, where action planning processes differ markedly. This finding also suggests that a late efference copy signal is critical; motor preparation takes longer and therefore starts earlier for antisaccades compared to prosaccades, but this early activity does not give rise to a larger chronostasis effect.

Although a subcortical signal may *trigger* chronostasis and directly initiate certain timing operations, any retrospective adjustment of perceptual content is presumably generated elsewhere. So where is this signal transmitted? We have speculated that the experience of saccadic chronostasis may reflect receptive field shifts of visual neurons. These were first described in the lateral intraparietal area (LIP) of behaving monkeys (Duhamel et al. 1992) and have been found more recently in a number of other brain areas (Walker et al. 1995; Umeno & Goldberg 1997; Nakamura & Colby 2002). Some apparently retinocentric cells in these areas begin to respond *before a saccade has been initiated* to stimuli at locations that the saccade will bring into their receptive fields. The timing of this presaccadic activity varies widely across cells, but a brain region capable of averaging these neurons' initial responses to a postsaccadic stimulus could contribute to the experience reported during saccadic chronostasis experiments. Although the idea that receptive fields shift in response to an efference copy signal from the superior colliculus is physiologically plausible (Sommer & Wurtz 2002), the part played by such cells in producing saccadic chronostasis remains hypothetical.

10.3 Methodological issues

10.3.1 Is saccadic chronostasis simply an order effect?

When two or more intervals are presented in sequence, participants often exhibit biases in their temporal judgments. The best known example is the time order error (see Hellstroem 1985 and Allan 1979 for reviews). Subjects' judgments are often biased such that two identical, consecutively presented intervals do not appear of equal duration. Either interval can appear prolonged, and the direction and magnitude of the bias is difficult to predict. There are other examples of specific biases arising as a result of sequential presentation of stimuli. Rose and Summers (1995) reported that, when four squares are presented with intervening blank periods, the first and the fourth square seem prolonged compared to the middle two. It is also possible to observe the influence of one interval on another when one of these intervals is evaluated with a comparison stimulus presented much later (e.g., Sasaki et al. 2002). However, although others suggest that chronostasis is caused by the same mechanism as sequence effects (Hunt et al. 2008), none of these biases is directly relevant to saccadic chronostasis. Chronostasis is always evaluated relative to a control condition with identical sequential properties. Demonstrations of saccadic chronostasis, therefore, reveal a bias in subjective duration *over and above* any order effects that may be present in the particular procedure employed.

10.3.2 Do constant fixation conditions provide a suitable control?

The purpose of the constant fixation conditions in saccadic chronostasis experiments is to provide a match for the pattern of visual stimulation experienced in saccadic conditions. Three different kinds of control conditions have been used. The first type matches sequence effects (see section 10.3.1) but provides only an approximate match for visual stimulation. For example, Yarrow et al. (2001) and Park et al. (2003) used a numeric counter ("0," "1," "2," "3," "4") in fixation conditions (judge the "1" relative to the "2" and the "3"). In saccade conditions, subjects fixated a cross, then saccaded to the same counter, which changed to display a "1" midsaccade then progressed through the same numerical sequence. Hence foveal stimulation differed somewhat between the two conditions. In saccade conditions, subjects foveated a cross, then had a brief period of smeared foveal input during the saccade itself, followed by foveation of the target stimulus ("1"). This was compared to control conditions in which they foveated a "0" immediately followed by a "1."

The second type of control condition better approximates foveal stimulation by matching the first foveal stimulus (usually a cross) and introducing a brief blank period between it and the target stimulus. The blank period is intended to approximate the time the eyes were in motion in saccade conditions. For example, in the experiments of Yarrow et al. (2006a,b), running averages were calculated for saccade duration, and these were used to make sure the blank period was precisely matched to the duration of the saccade. In fact, this level of precision is probably not required. Yarrow et al. (2004a) ran an experiment evaluating perceived duration in four variants of the standard control condition. The cross changed to the target stimulus either immediately, after 50 msec, after 100 msec, or after 500 msec. Duration estimates were similar in all conditions, so the presence of a gap doesn't seem to affect perceived duration (although it does affect temporal order judgments; Yarrow et al. 2006a). Overall, these sorts of control conditions do a reasonable job of matching *foveal* stimulation under the assumption of saccadic suppression but leave open the issue of whether the visual motion sensed during the saccade might yield a chronostasis effect.

A third type of control condition attempts to answer this concern by having the critical visual objects in the control condition move in a way that approximates their motion on the retina in the saccade condition. In a recent example, Yarrow et al. (2004a) (Experiment 3) had subjects fixate a cross, while a second cross was displayed 20 deg away on the screen. Both crosses were reduced in contrast and then moved with near saccadic velocity

(200 deg/sec) such that the second cross moved toward fixation and the first cross moved away from fixation in a consistent manner. Halfway through this movement, the second cross changed into the target stimulus (a circle). At the end of the movement, subjects were left fixating this circle (now at full contrast) and made a judgment about the duration for which they had fixated it. This condition was compared with two variants of the more typical control condition and yielded similar PSEs.

Taken collectively, these results make saccadic chronostasis arising from foveal visual factors unlikely. However, it is currently uncertain whether full-field visual motion exactly matching that occurring during a saccade could yield a chronostasis effect. For this reason, further experimentation is required. If stimuli were presented via a mirror that could be rapidly rotated, it would be possible to produce full-field motion with a saccadic time course (e.g., Diamond et al. 2000). Duration estimates could be assessed for a stimulus brought to fixation using this approach and compared with a matched saccadic condition so that chronostasis could be positively demonstrated over and above full visual field stimulation.

10.3.3 Is it really the first interval that is being affected?

The standard chronostasis procedure involves comparing one interval with one or more subsequent intervals. This procedure cannot distinguish between biases that affect the first interval and those that affect later intervals in the opposite direction. Our assertion that the first interval is subjectively lengthened is, however, supported by our results using a temporal order judgment procedure (Yarrow et al. 2006a). It is further supported by an experiment in which a different kind of duration judgment was required. Yarrow et al. (2006b, Experiment 5) presented only a single postsaccadic stimulus (in these experiments a moving circle) and had subjects make absolute duration estimates (in milliseconds, msec) to evaluate its perceived duration. As expected, estimates were higher in the saccade condition compared to the control condition.

10.3.4 Is saccadic chronostasis an artifact of correcting presentation times in order to calculate points of subjective equality relative to the moment of foveation?

In the standard saccadic chronostasis procedure, the PSEs reported in saccade conditions are not simply calculated using the duration for which the target stimulus appeared on the screen in each trial (see Fig. 10.1). These PSEs incorporate an additional correction to display times. The rationale for this correction is as follows. The target stimulus changes into its postsaccadic state during the saccade, at a time when perception is degraded (Ross et al. 2001). We have assumed that it is not perceived to a degree compatible with the initiation of a mental timing operation until it is actually foveated. Hence, the time for which the stimulus was on screen during the saccade (the period from stimulus change to saccade termination; D_S in Fig. 10.1) is subtracted from presentation times before PSEs are calculated. The effects we report (the difference between control and saccade PSEs) can therefore be broken into two components: an increase in perceived duration relative to the on-screen presentation time, and our correction.

If this correction is not justified, there are two implications. First, the magnitude of the saccadic chronostasis effect would be overestimated. Note, however, that in all saccadic chronostasis experiments reported to date, omitting the correction would not have eliminated or reversed the direction of the effect. Put another way, we always obtain an increase in perceived duration relative to on-screen presentation time, even before the correction is applied.

A second implication, however, is more critical for the antedating hypothesis. The finding that the magnitude of saccadic chronostasis increases with saccade duration provides an important foundation for this account. In the original experiment reporting this effect (Yarrow et al. 2001, Experiment 1), the change to the target stimulus was triggered based on a set proportion of the distance into the saccade. This means that the size of the correction varied in the short and long saccade conditions, being larger in the latter case. Hence, if the correction is unwarranted, the saccade size difference may be artificially enhanced.

Because of its importance for interpreting the saccade length effect, the appropriateness of the correction was tested by Yarrow et al. (2001, Experiment 1C). They compared two saccadic conditions, both of which employed a large eye movement. In one condition, the change to the target stimulus was triggered near the beginning of the saccade. In a second condition, it was triggered near the end of the saccade. Recall that our correction equals the interval from the change trigger to the end of the saccade. This means that the size of the correction was large in the first condition and small in the second condition. Consider first the hypothesis that subjects did not perceive the midsaccadic change of stimulus or were uncertain about its timing and antedated their subsequent percept to a moment just before saccade initiation regardless of this event. In this case, we would expect corrected PSEs to be identical in both conditions but uncorrected PSEs to vary by the same interval that separated the trigger times in the two conditions. Now consider the alternative hypothesis – that subjects perceived the midsaccadic change of stimulus and used it as the start point in estimating the duration of the postsaccadic stimulus, with chronostasis yielding some constant addition to this estimate. In this case, we would expect corrected PSEs to differ by an amount equal to the temporal separation between the two trigger times, but uncorrected PSEs should not differ.

In this experiment, the interval between trigger times was 85 msec. We originally reported *corrected* PSEs, which differed by only 11 msec. This difference in PSEs was not significant, supporting the antedating view. There is an interpretational issue here because the conclusion depends upon a negative result (power $= 0.71$ two tailed, 0.8 one tailed). A reanalysis of the data from this control experiment using *uncorrected* PSEs shows a significant 75-msec difference ($t = 2.0$, $df = 9$, one-tailed $p = 0.036$), thus providing more positive support for the antedating account.

Hunt et al. (2008) have recently challenged the validity of the correction procedure based on a different kind of experiment. Their subjects made a 25-deg saccade from a cross to a counter initially showing a “0.” The counter changed to a “1” midsaccade, but only after the brief (25-msec) presentation of either an "x" or a "+" at the same location. Subjects were asked to discriminate between these two symbols, and indeed were able to do so. Hunt et al. (2008) therefore conclude that, in saccadic chronostasis experiments, subjects are able to see the midsaccadic change to the target stimulus, and that the correction is therefore flawed, undermining the saccade size difference effect. We believe that their conclusion is unwarranted because Hunt et al.'s subjects were performing a very different task to the one typically required in chronostasis experiments. They were asked to discriminate a brief midsaccadic event rather than judge the duration of a postsaccadic stimulus. This difference implies attending to the stimuli in different ways. Furthermore, the stimuli they used were probably larger and therefore more easily discriminated than those used in our experiments (although it is difficult to be certain because they report the point size of the typeface rather than the visual angle subtended). The impact of the midsaccadic stimulus change is better assessed in the same context used to demonstrate chronostasis in the first place, as in the trigger time experiment reported previously. The question is not whether a midsaccadic stimulus change can be perceived. The question is whether it is used as a time marker in saccadic chronostasis experiments.

To determine conclusively whether the saccade size effect is real or an artifact, we have conducted an experiment comparing saccades of different sizes (Yarrow et al. 2006a). We introduced a critical procedural change in this experiment. Instead of triggering the change to the target stimulus a set proportion of the distance into the saccade, this change was triggered at a similar time relative to the *end* of the saccade. Hence for both long and short saccades, the change was triggered around 30 msec before the target was fixated. The correction applied to PSEs was therefore *virtually identical in both conditions*. We nonetheless obtained a significant difference between PSEs in long and short saccade conditions. This finding provides clear evidence for a saccade size effect in saccadic chronostasis that cannot be ascribed to our correction technique and accords with the antedating account.

10.3.5 Is saccadic chronostasis really a perceptual phenomenon?

Do we really see (or recall seeing) an extended interval following a saccade, or could saccadic chronostasis result from some kind of response bias? In most chronostasis experiments, subjects judge the first interval relative to subsequent intervals, so a simple bias to respond "longer" would yield reduced PSEs. However, saccadic chronostasis is measured relative to a control condition, so any bias would have to be specific to saccade conditions. Perhaps, then, the presence of a saccade biases subjects toward making a "longer" response for some nonperceptual reason? This is also unlikely because the effect has been demonstrated when judgments are made about whether the second interval is longer or shorter than the first (Yarrow et al. 2004a, Experiment 4). In this case, subjects tended to respond "shorter" with equal display durations. Our method, however, cannot be said to be "criterion-free" in the sense derived from signal detection theory. For example, it is possible that our subjects employed some sort of high-level reasoning strategy in reaching their decisions. Although we asked our subjects to judge how long they *saw* the postsaccadic stimulus, they might have reasoned that this stimulus appeared during their saccade. Hence the display may have given rise to demand characteristics that encouraged subjects to compensate for their own saccades. We cannot completely discount this possibility, but the differences we find for saccade extent imply that this strategy would have to be extremely sophisticated. Moreover, this account does not fit with the phenomenology of the task. In our experience as observers, there is no sense of adding time or interpreting one's eye movements, only of accurately reporting a percept.

10.4 Saccadic chronostasis and other temporal illusions

Since the term "chronostasis" was first coined to describe the saccadic illusion that forms the focus of this chapter, a number of other illusions have been described and explicitly related to the effect. Alexander et al. (2005; see also Hodinott-Hill et al. 2002) played subjects five tones bounding four consecutive intervals. The tones were either presented all to one ear (the control condition) or the first tone to one ear and all subsequent tones to the other ear. The first interval seemed prolonged in comparison to the subsequent intervals when the auditory stimuli bounding it were presented to different ears, with PSEs reduced by around 160 msec compared to the control condition. In a subsequent experiment, the tones were presented to the same ear in all conditions, but features of the two tones bounding the first interval were manipulated. Increasing the volume of the second tone yielded somewhat reduced PSEs (an effect of around 50 msec), but no effect was obtained when the volume of the first tone was increased or when the pitch of the second tone was changed.

Hodinott-Hill et al. (2002) and Alexander et al. (2005) argue that these effects are linked to saccadic chronostasis. They also consider the manual chronostasis effects described next, as well as the sequential effect reported by Rose and Summers (1995), in which the first and last stimuli in a sequence of four are overestimated to be members of the same family. They make this argument based on parsimony and a broad similarity in the direction and magnitude of these effects. They then go on to offer an explanation based on arousal and/or action increasing the rate of a hypothetical internal clock. We feel that the grouping together of these particular illusions and the exclusion of other temporal biases is arbitrary. At face value, the procedures used by Alexander et al. (2005) have more in common with those employed by Grondin and coworkers, who have shown temporal biases when visual bounding stimuli appear at different spatial locations (Grondin 1998) and when auditory bounding stimuli have different spectral properties (Grondin & Rousseau 1991), than they do with typical demonstrations of saccadic chronostasis. Furthermore, when we formulated predictions based on the arousal account and tested them experimentally, we found no evidence that it could explain the saccadic chronostasis illusion (Yarrow et al. 2004a; see Section 10.2.2). The effects we have obtained using temporal order judgments also seem inconsistent with it.

Other temporal illusions motivated by the saccadic chronostasis finding have also been described. Yarrow and Rothwell (2003) asked subjects to make reaching movements toward a vibrating tactile stimulus that marked out target and reference intervals. Subjects overestimated the duration of the postmovement interval by 60–120 msec compared to a static control condition, but the size of the effect did not change for reaches of different extents/durations. Yarrow and Rothwell (2003) also tested conditions in which subjects viewed a sequence of visual intervals initiated by reaching to and/or pressing a button. PSEs for the first interval did not differ between movement and control conditions in any of three experiments, with experimental powers ranging from 0.8 to 0.99. The contrast between reaching to a tactile stimulus and having a reach/button press trigger a visual stimulus led Yarrow and Rothwell (2003) to speculate that the tactile effect depended upon uncertainty regarding the physical onset of the target stimulus. Subjects could feel the vibrator as soon as they touched it but could not know exactly when it had been switched on. In visual conditions there was no uncertainty because the visual stimulus was constantly fixated. However, in a different experiment that maintained an even closer correspondence to the saccadic chronostasis procedure, the first interval in a visual sequence was not subjectively prolonged in a patient with congenital ophthalmoplegia, who made "saccadic head movements" to a digital counter (Jackson et al. 2005). In this case, uncertainty would seem to be present, but no effect was reported.

In contrast to Yarrow and Rothwell's (2003) results, other groups have obtained chronostasis-like effects when a button press initiates a sequence of visually defined intervals. Park et al. (2003) initiated a digit sequence either randomly, 500 msec after a key press, or immediately after a key press. PSEs for the first interval were reduced by around 70 msec when the key press initiated the sequence compared to the random and delayed conditions. Park et al. (2003) also observed a similar effect when a vocal signal initiated the digit sequence. In a different set of experiments, Hunt et al. (2008) found a key press effect of around 70 msec in two experiments using a digit sequence, but not in three similar subsequent experiments. The exact conditions that give rise to this effect have yet to be pinned down.

If many different kinds of movement give rise to chronostasis-like effects, what common mechanism might explain these results? Park et al. (2003) suggest that another temporal illusion occurring in the context of movements might be responsible for their findings. Haggard et al. (2002) have reported that, when a brief auditory stimulus is presented shortly after a button press and is contingent upon that action, the subjective estimate of the action's time of occurrence is shifted in the direction of the auditory stimulus. Furthermore, the subjective time of occurrence for the auditory stimulus is shifted in the direction of the action, prompting these authors to label this effect "intentional binding." Generalizing this result, it is possible that *any* action shifts the perceived time of its sensory consequences toward the action. When such sensory consequences mark the first boundary of an interval, we might therefore expect the perceived duration of that interval to be increased.

Are intentional binding and chronostasis related? One approach to answering this question would be to explore in detail the psychophysical properties of each illusion and see how they compare. For example, intentional binding arises only when actions are made voluntarily (Haggard et al. 2002). The effect is actually reversed when a movement is initiated in the absence of volition (by using transcranial magnetic stimulation over primary motor cortex). Saccadic chronostasis, however, occurs with a similar magnitude for voluntary and highly automatic classes of saccades. This dissociation implies that the two illusions may result from different kinds of mechanisms (Yarrow et al. 2004b). One could also ask whether other illusions share properties we have demonstrated for saccadic chronostasis, such as a constant effect size regardless of stimulus duration (Yarrow et al. 2004a) and a shift in the perceived time of the event that initiates the critical interval (Yarrow et al. 2006a). We already know, for example, that although saccadic chronostasis depends upon the duration of the preceding movement, tactile chronostasis does not (Yarrow & Rothwell 2003). Another possibility would be to assess the strength of each illusion using the same subject group and look for correlations. It is possible that there is a common mechanism underlying some or all of the biases described here, or that such a mechanism may be supplemented by more specific mechanisms in particular cases. However, at this point such links seem speculative. We prefer to view these psychophysical biases as separate until some convincing evidence emerges to link them. In our view, much of the confusion regarding common mechanisms in chronostasis arises from confusing action-specific effects with sequence position effects. We are not aware of any evidence implying that these kinds of effects are related.

10.5 Conclusions

When observers saccade toward a visual target, they overestimate the duration for which it is presented. We have conducted extensive investigations of this illusory bias and favor an antedating account in which the saccade target is subjectively experienced as having been fixated since before the eye movement began. This account explains why we have no temporal experience corresponding to the period of our saccades and therefore helps explain our conscious experience during active vision. Although a number of other illusory biases have been linked with saccadic chronostasis, their relationship to it remains currently unclear. Where direct evidence linking these effects has been sought, it has generally not been obtained. Hopefully, further research will allow the mechanisms underlying these various biases to be better understood.

References

Alexander, I., Thilo, K. V., Cowey, A., & Walsh, V. (2005). Chronostasis without voluntary action. *Experimental Brain Research* **161**: 125–132.

Allan, L. G. (1979). The perception of time. *Percept Psychophys* **26**: 340–354.

Brown, P., & Rothwell, J. C. E. (1997). Illusions of time. *Society for Neuroscience Abstracts, 27th Annual Meeting* **23**: 1119.

Campbell, F. W., & Wurtz, R. H. (1978). Saccadic omission: why we do not see a grey-out during a saccadic eye movement. *Vision Res* **18**: 1297–1303.

Diamond, M. R., Ross, J., & Morrone, M. C. (2000). Extraretinal control of saccadic suppression. *J Neurosci* **20**: 3449–3455.

Duhamel, J. R., Colby, C. L., & Goldberg, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science* **255**: 90–92.

Fischer, B., & Ramsperger, E. (1984). Human express saccades: extremely short reaction times of goal directed eye movements. *Experimental Brain Research* **57**: 191–195.

Grondin, S. (1998). Judgments of the duration of visually marked empty time intervals: linking perceived duration and sensitivity. *Percept Psychophys* **60**: 319–330.

Grondin, S., & Rousseau, R. (1991). Judging the relative duration of multimodal short empty time intervals. *Percept Psychophys* **49**: 245–256.

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nat Neurosci* **5**: 382–385.

Hellstroem, A. (1985). The time-order error and its relatives: mirrors of cognitive processes in comparing. *Psychol Bull* **97**: 35–61.

Hodinott-Hill, I., Thilo, K. V., Cowey, A., & Walsh, V. (2002). Auditory chronostasis: hanging on the telephone. *Curr Biol* **12**: 1779–1781.

Hopp, J. J., & Fuchs, A. F. (2002). Investigating the site of human saccadic adaptation with express and targeting saccades. *Experimental Brain Research* **144**: 538–548.

Hunt, A. R., Chapman, C. S., & Kingstone, A. (2008). Taking a long look at action and time perception. *J Exp Psychol Hum Percept Perform* **34**: 125–136.

Jackson, S. R., Newport, R., Osborne, F., Wakely, R., Smith, D., & Walsh, V. (2005). Saccade-contingent spatial and temporal errors are absent for saccadic head movements. *Cortex* **41**: 205–212.

Nakamura, K., & Colby, C. L. (2002). Updating of the visual representation in monkey striate and extrastriate cortex during saccades. *Proc Natl Acad Sci USA* **99**: 4026–4031.

Park, J., Schlag-Rey, M., & Schlag, J. (2003). Voluntary action expands perceived duration of its sensory consequence. *Experimental Brain Research* **149**: 527–529.

Rose, D., & Summers, J. (1995). Duration illusions in a train of visual stimuli. *Perception* **24**: 1177–1187.

Ross, J., Morrone, M. C., Goldberg, M. E., & Burr, D. C. (2001). Changes in visual perception at the time of saccades. *Trends Neurosci* **24**: 113–121.

Sasaki, T., Suetomi, D., Nakajima, Y., & ten Hoopen, G. (2002). Time-shrinking, its propagation, and Gestalt principles. *Perception and Psychophysics* **64**: 919–931.

Sommer, M. A., & Wurtz, R. H. (2002). A pathway in primate brain for internal monitoring of movements. *Science* **296**: 1480–1482.

Treisman, M., Faulkner, A., Naish, P. L., & Brogan, D. (1990). The internal clock: evidence for a temporal oscillator underlying time perception with some estimates of its characteristic frequency. *Perception* **19**: 705–743.

Umeno, M. M., & Goldberg, M. E. (1997). Spatial processing in the monkey frontal eye field. I. Predictive visual responses. *J Neurophysiol* **78**: 1373–1383.

Walker, M. F., Fitzgibbon, E. J., & Goldberg, M. E. (1995). Neurons in the monkey superior colliculus predict the visual result of impending saccadic eye movements. *J Neurophysiol* **73**: 1988–2003.

Wearden, J. H., Edwards, H., Fakhri, M., & Percival, A. (1998). Why "sounds are judged longer than lights": application of a model of the internal clock in humans. *Q J Exp Psychol B: Comparative and Physiological Psychology* **51**: 97–120.

Yang, Z., & Purves, D. (2003). A statistical explanation of visual space. *Nat Neurosci* **6**: 632–640.

Yarrow, K., Haggard, P., Heal, R., Brown, P., & Rothwell, J. C. E. (2001). Illusory perceptions of space and time preserve cross-saccadic perceptual continuity. *Nature* **414**: 302–305.

Yarrow, K., Haggard, P., & Rothwell, J. C. (2004a). Action, arousal, and subjective time. *Conscious Cogn* **13**: 373–390.

Yarrow, K., Johnson, H., Haggard, P., & Rothwell, J. C. E. (2004b). Consistent chronostasis effects across saccade categories imply a subcortical efferent trigger. *J Cogn Neurosci* **16**: 839–847.

Yarrow, K., & Rothwell, J. C. E. (2003). Manual chronostasis: tactile perception precedes physical contact. *Curr Biol* **13**: 1134–1139.

Yarrow, K., Whiteley, L., Haggard, P., & Rothwell, J. C. (2006a). Biases in the perceived timing of perisaccadic visual and motor events. *Perception and Psychophysics* **68**: 1217–1226.

Yarrow, K., Whiteley, L., Rothwell, J. C., & Haggard, P. (2006b). Spatial consequences of bridging the saccadic gap. *Vision Res* **46**: 545–555.

11

Experiencing the future: the influence of self-initiation on temporal perception

TIMOTHY VERSTYNEN, MICHAEL OLIVER, AND RICHARD B. IVRY

Summary

Anticipation is a hallmark of skilled movements. For example, when removing plates from a loaded tray, the upward force generated by the supporting hand is reduced in anticipation of the reduced load. An adjustment of the postural force occurs as a result of the predicted consequences of the self-initiated action. Although the effect of anticipatory processes is easily discerned in the actions themselves, it is unclear whether these processes also affect our perceptual experience. In this chapter we focus on the relationship between action and the perceptual experience. We begin by reviewing how actions provide reliable predictions of forthcoming sensory information. Following this, we discuss how the anticipation of the time of external events is an important component of action-linked expectations. Finally, we report two experiments that examine how temporal predictions are integrated with the incoming sensory information, evaluating whether this integration occurs in a statistically optimal manner. This predictive process provides the important advantage of compensating for lags in conduction time between peripheral input and the central integration of this information, thus overcoming the physical limitations of sensory channels.

11.1 Racing against sensory delays

An important problem for the brain to solve is how to compensate for the temporal gap between when a stimulus is registered by a sensory detector and when it is recognized, either consciously or subconsciously, in the cortex. In humans, such delays happen on the order of hundreds of milliseconds (for review, see Welch & Warren 1986). This slow conduction time can greatly impair how we react to events in our local environment. In many cases, this can mean the difference between smooth, coordinated actions and clumsy gestures. For example, reaching movements would be inherently unstable if the final position were achieved solely on the basis of visual and proprioceptive feedback signals. Such delays would invariably lead to significant endpoint oscillations (Lacquaniti et al. 1982; Soechting et al. 1986).

The negative effects of sensory lags can be attenuated by maximizing the efficiency of the input channels. Such an approach has been applied in the field of robotics with great

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

success (Blomdell et al. 2005). Unfortunately, the physiology of the human nervous system limits the speed of conduction of sensory signals. Therefore, the brain must estimate the true instantaneous state of the world by adopting a strategy of predictive inference based on existing sensory information. This has several advantages. First, perceptual processing can be enhanced by focusing attention to environmentally relevant events before they happen. Second, it can increase the reactive speed of our actions by compensating for the delays between the activation of sensory receptors and central processing.

There are many ways by which the brain extrapolates from current information to anticipate future sensory events. *Spatiotemporal regularities* are exploited to predict when and where a stimulus will occur. For example, neurons in the primary visual cortex will respond more vigorously to a visual stimulus that is repeatedly presented in the same part of the visual field and at predictable intervals than to a stimulus that randomly appears in its receptive field (Sharma et al. 2003). *Motion extrapolation* is another method by which we predict the future location of a moving object. Such processes need not be veridical. For example, a moving stimulus is perceived as being further along its trajectory when compared to a stationary object, even if the two objects are at the same location (Nijhawan 1994). We also make use of *general physical principles or intuitions*, based on our experience of consistent environmental features. For example, because larger objects are generally heavier than smaller objects we make *a priori* inferences about the weight of an object based on its size. This prediction also can distort our perception, as in the size–weight illusion (Charpentier 1891), whereby the smaller of two objects of equal weight is perceived as the heavier of the two.

11.2 The influence of action on perception

Action-linked expectations manifest themselves in many of our movements. Anticipatory postural adjustments constitute one class of such actions. Consider the waiter arriving at a table of diners with a tray full of entrees. His ability to keep the tray steady as he serves each plate is quite impressive given that more than 100 msec of processing delays can be expected before his brain will register the change in the downward force that occurs as each plate is lifted. Stability is maintained by a predictive reduction of the upward force generated by the supporting hand against the tray. This reduction must occur just before each successive plate is removed by the active hand (Hugon et al. 1982; Dufossae et al. 1985; Lum et al. 1992; Masson et al. 1999). These anticipatory predictions appear to be intimately linked to the waiter's self-generated actions, as they are not observed when the cues for the unloading are only sensory in nature (Diedrichsen et al. 2003). If a helpful patron assists by reaching over and taking her own entrée, an unfortunate mess may occur. If the resistive force generated by the supporting hand remains too large, the resulting upward displacement of the tray may launch the remaining plates into their individual orbits.

Does the anticipatory response of the waiter also influence his perceptual experience? The many recursive connections between motor and sensory areas in the brain suggest a relationship of mutual influence (for review, see Nelson 1996). Indeed, there are many demonstrations of situations in which our actions influence our perceptions. A classic example involves the “tickling effect” (Weiskrantz 1971). If someone brushes the tip of their finger across the palm of your hand, the tactile sensation can be quite ticklish. However, when you use your other hand to create a near-identical tactile stimulus, the ticklish sensation is attenuated. Thus, the sensory experience is more intense when it results from the actions of an external agent compared to when it is self-produced. The attenuated experience from self-stimulation results from our ability to precisely predict, in space and time, the tactile stimulus (Blakemore et al. 1999). In the self-stimulation condition, the insertion of a slight delay between the participant’s own movement and the resulting tactile stimulus will lead to an increase in the intensity of the tactile experience. Altering the angle between the direction of the movement and the tactile stimulus will also increase the perceived intensity. Thus, the attenuation of ticklishness is maximized when the somatosensory experience on the palm of the hand matches the spatiotemporal profile of the action used to produce it.

Another powerful example of how action-linked predictions influence perception is the occuloparalytic illusion, described by Matin and colleagues (1982). After receiving injections of d-turbocurarinem, a cholinergic antagonist of the muscle spindles and extrafusal muscles that leads to the transient paralysis of eye movements, participants were asked to localize visual and auditory targets in space. Perceived location of external targets was greatly influenced by how the participants intended to move their eyes. Thus, action-based changes in visual perception can result from anticipatory consequences of motor commands to the eyes, even if those actions are not implemented.

How might a phenomenon such as the tickling effect or the occuloparalytic illusion extend to the sensory experience of our hypothetical waiter as he serves the dinner entrees? Is his perception of the tray’s weight affected by his actions? Predictive mechanisms allow for anticipatory postural adjustments prior to the volitional actions required to remove each plate. But does this expectation also modulate his perceptual experience of the force imposed by the tray?

To examine this issue, we tested a group of participants in a modified version of the unloading task (Diedrichsen et al. 2007). Participants were situated in a 3D visual-haptic virtual reality environment where they could manipulate simulated objects. Participants were instructed to lift and hold a platform-like object with one hand (the postural hand; Fig. 11.1(a)). After a brief stabilization period, the virtual object was lifted, either by an action of the participant, *self-unloading*, or by an unseen agent, *external-unloading*. During self-unloading, a cue instructed the participant to use his or her other hand to lift the object as quickly as possible. On external-unloading trials, the forces acting on the object were programmed such that the object rose off the supporting hand in a manner that simulated the dynamics of self-unloading; the participant did not move the other hand on these trials.

As noted previously, self-unloading results in a reduced upward perturbation of the postural hand compared to external-unloading. This effect could occur because of a well-timed reduction in the upward force generated by the postural hand or by increasing the stiffness of the postural hand, perhaps by cocontraction of antagonist muscles. The virtual

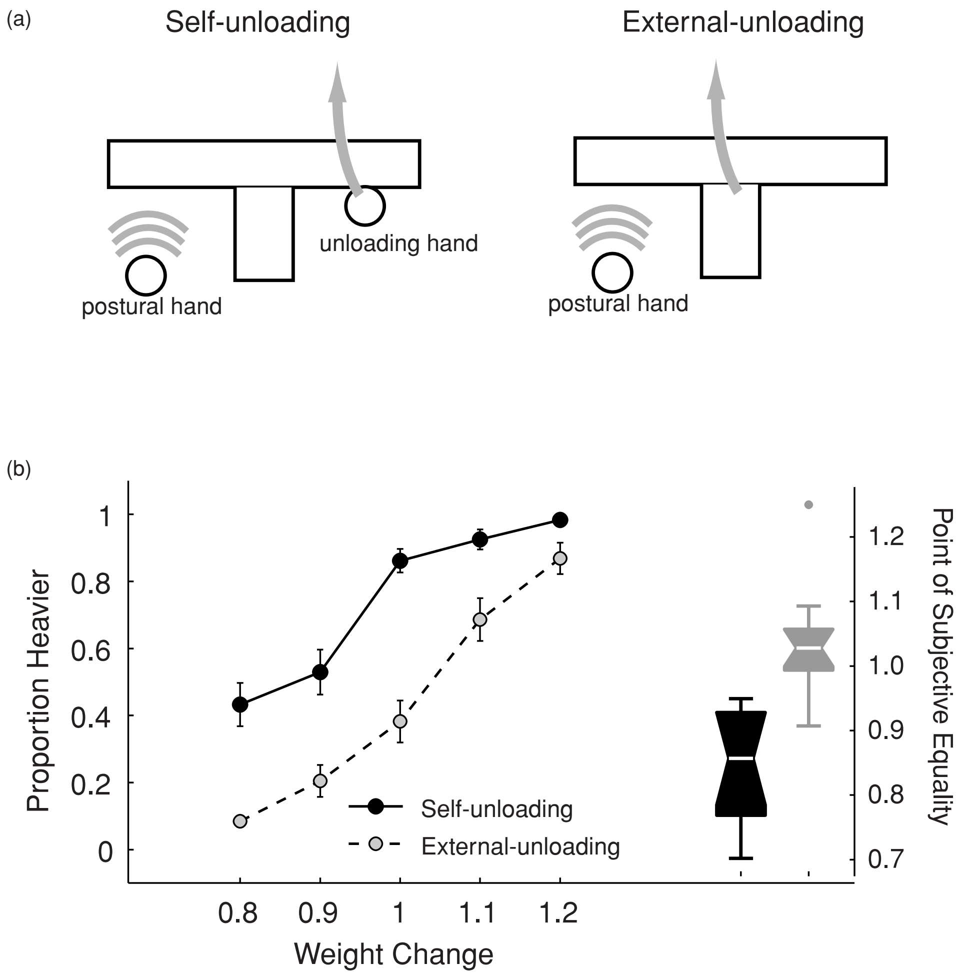


Fig. 11.1 Influence of action on force perception. (a) Participants were instructed to lift and hold a virtual object with one hand. Following a cue, participants had to either unload the object using the other hand (self-unloading) or observe the computer perform the lifting action (computer-unloading). On most trials, a "phantom" force that remained on the loaded hand was adjusted at the time of unloading to be heavier or lighter than the force of the object prior to unloading. (b) Psychometric functions of perceived force changes and point of subjective equality (PSE) estimates. Participants perceived the phantom force as being heavier during the self-unloading trials than during the computer-unloading trials. (Reprinted from Diedrichsen et al. 2007.)

reality environment can be exploited to contrast these two mechanisms by the inclusion of "catch" trials (Diedrichsen et al. 2003, 2005). On some of the self-unloading trials, the visual feedback, as well as haptic experience of the unloading hand, cues the participant that the object is being lifted from the postural hand. However, the downward force generated by the virtual object remains on the postural hand. A downward deviation of the postural hand is observed during these catch trials, indicating that the anticipatory response is in fact due to a reduction in the upward force rather than an increase in stiffness.

The "phantom" force used to create the catch trials can also provide a probe on whether perception is altered by the anticipatory mechanisms. Following self-unloading trials, we asked participants to judge the haptic experience of the postural hand. The initial force of the phantom objects was 3.5 g. Participants were told to report whether a force remained on the postural hand (as would be true on catch trials) or was absent (as would be true on natural unloading trials). On the catch trials, we altered the force of the object right at the time of unloading, with the resulting force ranging from 80–120% of the original 3.5 g force. By varying the phantom force, we could ask participants to not only report whether it was present or absent, but also whether, when present, it had "increased" or "decreased" in comparison to the force experienced prior to unloading.[1] For comparison, we also included external-unloading trials. Here, the object was seen to rise off the postural hand and, with similar probabilities, the force on that hand was either removed or persisted in a similar manner as on self-unloading trials.

Figure 11.1(b) shows the psychometric functions for the self- and external-unloading trials. On catch trials, participants consistently perceived the phantom force as greater during self-unloading trials compared to external-unloading trials. The point of subjective equality (PSE), defined as the force at which a participant is equally likely to judge the force as "increased" or "decreased," was significantly different between the two conditions ($t(11) = 5.51$, $p < 0.001$). Moreover, during self-unloading, the participants were biased to perceive the phantom force as having increased in comparison to the force experienced prior to unloading. On trials in which the force was unchanged, participants reported an increase of force on approximately 85% of the trials. We assume that the bias on self-unloading trials reflects a comparison process in which the sensory experience is compared to a representation that incorporates anticipated changes in the sensory experience due to self-generated actions.

This experiment suggests that our hypothetical waiter does, in fact, perceive the changes in the weight of the tray differently depending on whether he removes the plates himself or whether he is assisted by the eager patron. Whether reconstructing the visual environment (Matin et al. 1982), experiencing a ticklish tactile stimulus (Weiskrantz 1971; Blakemore et al. 1999), or perceiving haptic forces (Diedrichsen et al. 2007), our actions lead us to predict forthcoming changes in sensory information and influence our internal perceptual reconstruction of the environment.

11.3 Predicting the future

Predictive mechanisms anticipate sensory events and can compensate for delays in the conduction of sensory information from the periphery to the brain. This not only helps

---

[1] Although our main interest is in the perceived force on catch trials, it was necessary to include natural unloading trials because the anticipatory postural adjustment is rapidly extinguished by the catch trials (Diedrichsen et al. 2005). Thus, catch trials only occurred on 33% of the trials.

to construct veridical percepts but facilitates stable and appropriate actions as we interact with the environment. Given that these internal expectations are, at least in part, produced out of a necessity to overcome temporal delays in the processing of sensory information, action-induced expectations may also modulate our perception of the timing of sensory events.

Indeed, Yarrow and colleagues (2001) reported evidence that actions can directly modify our perception of time. They set out to investigate a puzzling illusion that has been noted since the advent of the analog clock; namely, why does the initial tick of the seconds hand appear to move slower when first glancing at a clock than subsequent ticks? In the extreme, it is frequently reported that the clock appears to briefly stop after the first tick, an illusion referred to as chronostasis. This phenomenon is discussed in detail in Chapter 10, so here we will provide only a brief summary of the Yarrow study. Participants were required to saccade to a stimulus, the number "0," that was positioned either 22 or 55 deg from fixation. Once the saccade was initiated, the stimulus changed to the number "1." This number was present for a variable duration, and subsequently incremented by one every 1000 msec. Thus, the perception was of a seconds counter incrementing periodically and participants were asked to judge if the duration of the initial number was shorter or longer than the other numbers. Participants judged the duration of the "1" to be longer than that of the subsequent numbers, and this effect was influenced by the amplitude (and thus duration) of the saccade. In fact, the temporal distortion was so profound that participants sometimes judged that the onset of the "1" occurred prior to saccade, which triggered the change.

This result led the authors to hypothesize that chronostasis arises from the need to "fill in" information that is lost during the saccade. Because the movement triggers the stimulus change, the estimate of the stimulus onset gets linked to the motor command that generates the saccade. A key assumption here is that the visual system is blind to visual information during the eye movement. Although it remains to be seen whether visual information is truly lost during saccades, two pieces of evidence run contradictory to this lost-time hypothesis. First, the chronostasis effect is not limited to the visual modality, but is also present for auditory (Hodinott-Hill et al. 2002) and tactile events (Yarrow and Rothwell 2003). It is highly unlikely that all sensory modalities are "blinded" during saccadic eye movements. Second, chronostasis is also observed when the trigger for the stimulus change is a key press rather than a saccade (Park et al. 2003).

An alternative explanation for this modulation of temporal perception is that the brain links the motor command instigating the action with the onset of the resulting sensory stimulus. This hypothesis was suggested in a related set of elegant experiments conducted by Haggard and colleagues (2002). Participants were first asked to time the occurrence of several events by observing the rotating seconds hand on a clock. In the single-event case, participants were instructed to report the time of either a voluntary key press, an involuntary muscle twitch induced by transcranial magnetic stimulation of the motor cortex, or an external auditory click. Temporal judgments were veridical in all three conditions. The participants were then tested in an "operant" condition in which a tone was presented

250 msec after each of the three events. The tone was perceived as occurring much earlier in the voluntary key press condition compared to the other two conditions. The authors suggest that the shift in the perceived timing of the tone is due to an association between the awareness of the voluntary action and its sensory consequence.

There is, however, one methodological concern with these studies. As noted previously, a moving object is perceived as being slightly ahead of its true state (Nijhawan 1994). This suggests that using the state of a moving clock as the dependent measure for time perception may introduce biases in estimating the timing of an event. Although it remains clear that temporal perception is influenced by self-generated actions differently than by other events (because the moving clock is the same for all conditions), the nature of the distortion is unclear because the dependent variable may reflect multiple processes. Moreover, the process by which a motor command influences the end sensory percept may be influenced by the attentional state of the participant in a different way than that which occurs in a purely perceptual condition. Spatial attention can alter the perceived temporal order of events as exemplified by the "shooting line" illusion (for review, see Schmidt 2000). It is possible that, when initiating an action, attention is briefly shifted to the appropriate area of space for the sensory consequence of that action. This shift may prime or alter the perceived onset of the sensory stimulus. Regardless of the underlying mechanism, the common finding from both the Haggard et al. (2002) and chronostasis studies is that voluntary actions appear to link the *onset* of the sensory percept with the motor commands used to instigate the actions themselves.

To further examine how the passage of time may be influenced by self-generated actions, we performed a series of experiments in which temporal judgments were made about static sensory events (Oliver et al. 2003). In the first experiment, eight participants were asked to compare the duration of two successive vibrotactile stimuli (800 Hz) presented to the middle finger of the left hand. The duration of one stimulus, the *standard*, was always fixed at 800 msec. The duration of the other stimulus, the *comparison*, varied from 600 msec to 1000 msec. The order of the standard and comparison was randomized across trials. Following the second stimulus, the participant indicated which stimulus (the first or second) was longer in duration. The important manipulation involved the manner in which the stimuli were initiated. On each trial, the participant would initiate one stimulus by pressing a response key with the middle finger of the right hand. The onset of the other stimulus was controlled by the computer. The self-initiated stimulus could either be the standard or comparison stimulus, counterbalanced for the two orders. The computer presented a visual cue prior to either stimulus indicating whether a key press was required. For computer-triggered stimuli, the cue "Computer" was displayed and the vibrotactile stimulus began after a variable delay of 200–500 msec. For self-triggered stimuli, the cue "Self" was displayed. The vibrotactile stimulus began immediately after the subsequent key press.

The results from previously described studies suggest that voluntarily initiated sensory events are perceived as starting earlier than their true, veridical onsets. From this we predicted that the self-initiated stimuli would be perceived as having a longer duration than computer-initiated stimuli. As shown in Fig. 11.2, the results are consistent with this prediction. For a given duration, participants were more likely to judge the comparison

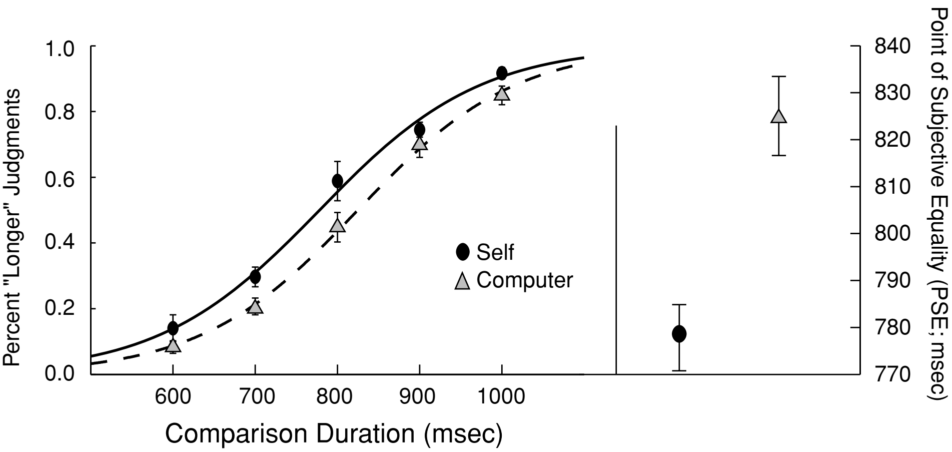


Fig. 11.2 Influence of self-initiation on vibrotactile perception. In a two-alternative forced choice experiment, participants were asked to compare the duration of a standard vibrotactile stimulus (800 msec) to that of a comparison stimulus ranging in duration from 600 to 1000 msec in steps of 100 msec. The stimulus was either initiated via a key press (self; black) or by an external agent (computer; gray). Participants consistently judged self-initiated comparison stimuli as being longer than computer-initiated stimuli of equal duration.

stimulus as longer for the self-initiated trials. This distortion is highlighted for trials in which the standard and comparison were both 800 msec. The self-initiated stimulus was judged "long" on approximately 60% of the trials when the two stimuli were, in fact, of equal duration.

To quantify this effect, we estimated the parameters of the psychometric function using binomial logistic regression

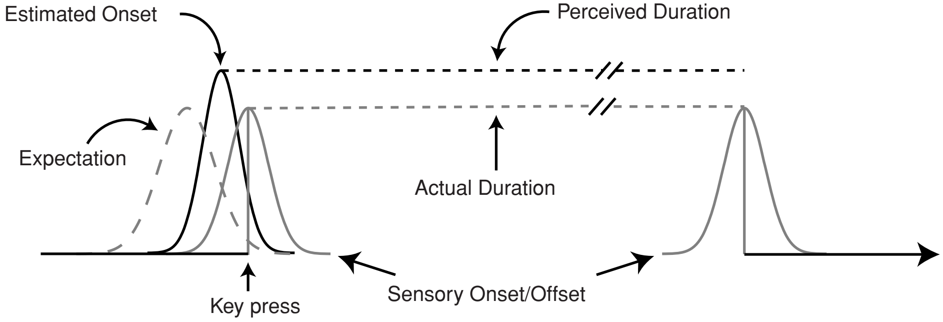
$$y = \lambda_D x_D + \lambda_0,$$

where "y" is the participant's binomial response on a given trial as to whether the comparison was longer or shorter than the standard and "x" is the duration of the comparison stimulus on that trial. The free parameters λ_D and λ_0 represent the weighted influence of changing stimulus duration on "y" and the subject's own baseline response bias respectively. Assuming that the curves follow a logistic probability function, the point of subjective equality (PSE) is calculated as

$$\text{PSE} = -1 * \lambda_0 / \lambda_D.$$

If participants underestimate the perceived duration of the comparison stimulus, then the PSE should be greater than the actual standard duration (describing a psychometric function shifted to the right). In contrast, if participants overestimate the comparison's duration, the PSE should be smaller than the standard duration (e.g., shifted function to the left). The mean PSE value for trials in which the self-initiated stimulus was the comparison is 23 msec less compared to when the self-initiated stimulus was the standard.

(a) No Delay Condition



(b) 50 msec Delay Condition

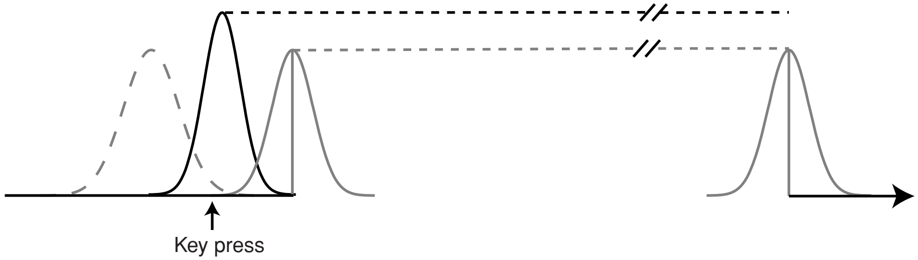


Fig. 11.3 Model in which self-initiated stimuli are perceived as being slightly longer than computer-initiated stimuli because of the incorporation of an a priori expectation generated by the motor command to trigger the stimulus. (a) Even when the stimulus is triggered immediately by the key press, the interval is perceived as longer because the perceived onset is a combination of the true sensory information and the action-linked expectation. (b) When a delay is inserted between the key press and the onset of the stimulus, the perceived onset (P[onset]) is shifted farther away from the true onset (P[haptic]). This is due to the influence of the action-linked expectation (P[expect]), which remains locked to the motor command used to produce the key press.

To summarize, when a tactile stimulus on one hand is triggered by an action performed by the other, the perceived duration of that stimulus is lengthened compared to *when the exact same stimulus is initiated by an external agent*. Unlike the temporal distortion studies reported previously, participants in our study did not have to refer to an external clock; the judgments were based on internal mechanisms for representing temporal information. Nonetheless, the results are quite similar. It appears that the percept of the onset of a stimulus is attracted to a movement that initiated the stimulus. More importantly, the shift in perceived onset is pulled toward the point at which the *intention to move* occurs: that is, well before the action is initiated.

We propose that information linked to the motor command, for example, an efference copy signal or forward model, is used to anticipate the perceived onset of the stimulus. Figure 11.3(a) presents a schematic diagram of this hypothesis. Central to this model is

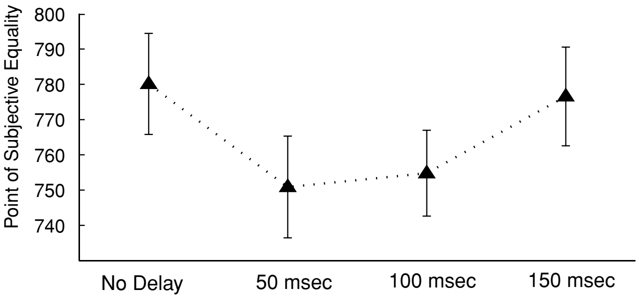


Fig. 11.4 Delay effects. Participants performed a temporal bisection task in which the onsets of all of the stimuli were triggered by a key press. Introducing a 50-msec delay between the key press and the stimulus onset resulted in an additional 30-msec increase in the perceived duration of the stimulus. For longer delays, the illusion was reduced.

the idea that the perceived onsets and offsets of a stimulus (e.g., the vibrotactile stimulus in our study) are not veridical. Noise is introduced not only because of delays due to transmission along sensory channels but also because of variability in the transmission itself, decay of the signals within these sensory channels, and noise in the central decision processes. Self-initiated stimuli have an advantage in that the movement command help create an expectation of the forthcoming stimulus. The percept of the stimulus onset is thus a combination of *a priori* information linked to the motor command and the sensory information coming from the periphery. In contrast, there is no motor signal for computer-initiated stimuli to supplement the expected onset time. The expected probability function is simply a uniform distribution prior around the general time of the onset of the stimulus, because any point in time is equally as likely to be the onset as any other time around the key press.

We will further develop this model in the next section. First, we wish to explore one prediction derived from the basic ideas of the model. Introducing a small delay between the key press and the onset of the stimulus should lengthen the perceived duration of self-initiated stimuli. We tested this prediction in another group of ten participants using a temporal bisection task. An exposure phase was first run in which participants were presented examples of short (600 msec) and long (1000 msec) vibrotactile stimuli. Each example was self-initiated by a key press and followed by feedback indicating whether the stimulus had been short or long. Following this exposure phase, participants completed a series of self-initiated trials in which a single stimulus was presented for a variable duration and the participant categorized it as short or long (without feedback). There were four conditions based on the interval between the time of the keypress and the onset of the vibrotactile stimulus. In the 0 msec, no-delay condition, the vibration started as soon as the key press was detected. In the other three conditions, a delay of either 50, 100, or 150 msec was inserted between the keypress and vibration onset (Fig. 11.3(b)).

Figure 11.4 shows the results for this experiment. The PSE for the no-delay condition is nearly identical to that observed in the self-initiation condition in our first vibrotactile experiment. Thus, the perceived duration of a self-initiated stimulus is lengthened even in comparison to an inferred standard duration (although this could also reflect a response bias in the second experiment). If this distortion is linked to the motor command, then we should observe that the effect will be enhanced when delays are introduced between the key press and stimulus onset. The results provide qualified support for this prediction. There is a 30-msec additional increase in the perceived duration during the 50-msec delay condition (evident by the shift of the PSE in the opposite direction), indicating that the estimate of stimulus onset is linked to a motor command triggering the stimulus. However, this effect is limited (see also Blakemore et al. 1999). There was no additional increase in perceived duration for the 100-msec delay condition and for the 150-msec delay condition, and perceived duration was similar to that observed in the no-delay condition. This nonmonotonicity suggests that when the delay becomes substantial, the contribution of the action-linked expectation is reduced and perceived duration is solely based on the stimulus.

In this section we discussed evidence indicating that such action commands may modify the perceived timing of sensory events. This influence on timing could be due to one of three nonexclusive mechanisms. First, an action may influence the perceived onset of the stimulus by altering the operation of an internal clock that marks the timing between events. Second, initiating an action may entail shifts in attention, and this could influence the perceived onset of a stimulus. Third, actions may modulate the perceptual processes, perhaps through the priming of perceptual channels by the motor commands, and thus advancing the time at which the stimulus onset is detected.

The model outlined in Fig. 11.3 does not differentiate between these hypotheses. Although all three mechanisms may be at play, there is some evidence to rule out the hypothesis that actions directly influence the perception of time. If action-linked changes in temporal perception were attributed to an adjustment in the rate of an internal clock, then the magnitude of this effect should also be affected by the duration of the stimulus being judged. However, Yarrow and colleagues have demonstrated that the magnitude of the chronostasis effect remains constant across a range of target stimulus durations (Yarrow et al. 2004). Thus, it does not appear that actions directly affect the internal representation of time itself.

We know from electrophysiological studies of the cat primary visual cortex that movements can directly influence principal sensory regions. The firing rates of V1 neurons increase during eye movements, even when these movements are made in the dark (Toyama et al. 1984). This saccade-related activity is also observed when the eye muscles are paralyzed, indicating that the V1 modulation is linked to the command to move the eyes rather than to the eye movements themselves. If we assume that an increase in baseline firing rate would lead to faster detection by bringing signal-induced firing closer to a decision threshold, then the perceived onset of a stimulus would occur earlier. This priming could be due to attentional shifts related to the movement command or direct corollary discharge from upstream motor regions.

11.4 Integrating expectations and sensory input

The preceding sections indicate that information linked to motor commands can influence perception of environmental events. We now turn to the question of how action- and sensory-based information is integrated.

There is emerging evidence that information *between* sensory modalities is integrated in a statistically optimal fashion (van Beers et al. 1999; Ernst & Banks 2002; Kording & Wolpert 2004). This integration process is unlikely to be limited to sensory information; internal, *a priori* expectations provide another salient source of information that should be combined with sensory signals (Weiss et al. 2002). It is reasonable to assume that the internal expectations linked to motor commands are combined with the sensory information in a similar, statistically optimal manner. In the remainder of this chapter we consider how this idea may help explain how a self-initiated movement influences the perceived duration of a stimulus triggered by that movement.

First let us assume that the perceived onset, P(onset) in Fig. 11.3(b), is the weighted contribution of sensory signals from the haptic system, P(haptic), and motor-linked expectations, P(expect). Thus we can define the perceived onset as

$$\mathrm{P(onset)} = \omega_{\mathrm{H}}\mu_{\mathrm{E}} + \omega_{\mathrm{E}}\mu_{\mathrm{H}},$$

where μ_{E} and μ_{H} represent the mean signals coming from the expectation and haptic channels, respectively. For the sake of simplicity we will assume that the mean of the expectation component occurs when the motor command is generated by the motor cortex and that the mean of the haptic component occurs when the sensory signal from the tactile stimulus is delivered to the sensory cortex in the brain.[2] For a no-delay condition, the difference between μ_{E} and μ_{H} represents the delay introduced between the activation of motor cortex until the tactile information is registered. The terms ω_{H} and ω_{H} represent the weighted influence of each of these distributions on the resulting perceived onset. The relative weighted influence of each channel should be proportional to the overall variance such that

$$\omega_{\mathrm{E}} = (1/\sigma_{\mathrm{E}}^2)/(1/\sigma_{\mathrm{E}}^2 + 1/\sigma_{\mathrm{H}}^2)$$
$$\omega_{\mathrm{H}} = (1/\sigma_{\mathrm{H}}^2)/(1/\sigma_{\mathrm{E}}^2 + 1/\sigma_{\mathrm{H}}^2).$$

Thus, the lower the variability of a particular channel (i.e., the more reliable that channel's information capacity is), the stronger its influence will be on the resulting percept.

Using the observations from the experiment in which we varied the delay between the key press and stimulus, we now show how the influence of the motor-linked expectation on perception (ω_{E}) can be estimated. To do this, we first want to understand how the perceived

---

[2] Which is equivalent to the time that the tactile stimulus is applied to the hand plus a constant delay in conduction time to the cortex. Because the delay would be equal for sensing both the onset and offset, this has little effect on the dynamics of our model.

onset (relative to the key press) changes as a function of the delay. We define this as

$$\Delta \mathrm{P(onset)} = \omega_{\mathrm{H}}\mu_{\mathrm{E}} + \omega_{\mathrm{E}}\mu_{\mathrm{Hdelay}} - (\omega_{\mathrm{H}}\mu_{\mathrm{E}} + \omega_{\mathrm{E}}\mu_{\mathrm{H0}}),$$

where μ_{Hdelay} and μ_{H0} reflect the mean onsets estimated from the haptic sensory channel during the delay and no-delay conditions, respectively. Thus, ΔP(onset) can simply be explained as

$$\Delta \mathrm{P(onset)} = \omega_{\mathrm{E}}(\mu_{\mathrm{Hdelay}} - \mu_{\mathrm{H0}}).$$

From this, the weighted influence of the motor-linked expectation can be defined as

$$\omega_{\mathrm{E}} = \Delta \mathrm{P(onset)}/\Delta\mu_{\mathrm{H}},$$

where $\Delta\mu_{\mathrm{H}} = \mu_{\mathrm{Hdelay}} - \mu_{\mathrm{H0}}$. For example, in the 50-msec delay condition, $\Delta\mu_{\mathrm{H}} = 50$. We now have a simple formula to estimate the influence of the internal expectation on perceived stimulus onset.[3] Using the observed values in the delay experiment $\omega_{\mathrm{E}} = 0.29$ when averaging across the three delay conditions, with considerable variation across the delay conditions (50 msec = 0.58, 100 msec = 0.25, 150 msec = 0.02). By definition that average weight term for the sensory input is $\omega_{\mathrm{H}} = 1 - \omega_{\mathrm{E}} = 0.71$. Thus, from this simplified example, it appears that the perceived onset is influenced to a greater degree by the actual onset of the tactile stimulus than by the movement-generated internal expectation of the onset.

One aspect of our results is at odds with the basic assumptions of this simple model. Contrary to what would be expected if the variance of the sensory channels and expectations were stationary, we found that perceived duration did not increase monotonically; there was no change for the 100-msec delay in comparison to the 50-msec delay, and it became shorter for the 150-msec delay. Within the framework of our model, this nonmonotonic pattern would reflect a change in the relative variances in either the haptic sensory channel or the motor-linked expectation. Assuming that the variance of the sensory channel remains constant over the different delays, at least for the range of values tested here, we would have to assume that the variance of the internal expectation increases for the longer delay conditions. This has the effect of "discounting" the internal expectation when there is a large mismatch between the timing of the expected and actual events. In the extreme, we would not expect any influence of the action if the stimulus onset occurred at an irregular time many seconds after the action. However, because the different delay conditions were randomly presented within the same blocks of trials, this effect cannot be explained by exposure to an "unreliable" expectation during longer delay trials.

There are several alternative accounts for the nonmonotonicity in the perceived duration function. First it could be that the integration of the sensory information and the expectation is a post hoc calculation, where the reliability of the expectation is revaluated prior to the generation of the final percept. This, however, fails to explain the electrophysiological

---

[3] In this simplified model, we assume that the mean offset from the haptic channel is veridical.

evidence that motor commands immediately influence primary sensory regions (Toyama et al. 1984) and would preclude a parsimonious account across the delay conditions. Another possibility is that the characteristics of the motor commands varied across the delay conditions. For example, the participants might have adopted a strategy of pressing the key until the vibration was felt despite our instructions to press the key in a quick and ballistic manner. Longer delays would result in more sustained presses, and this could influence the variance of the expectation generated from the motor command (e.g., if the variance of the expectation is linked to the velocity of the finger movement or the acceleration of the force of the key press). We did not measure key press duration in the current experiment and thus are unable to evaluate this hypothesis.

It is also possible that the perceived onset is not just a simple linear summation of an expectation and the sensory information but that this relationship is gain-modulated by a supervisory process that monitors the mean discrepancy between the two distributions (Blakemore et al. 2001). This would require the addition of a third parameter that represents the mean difference between the expected and sensed onset distributions. This new term would negatively weight the influence of the expectation distribution when there is a large discrepancy. Although such a model would more elegantly explain the data, it requires a supervisory process that monitors these discrepancies. Finally, it is entirely possible that the expectations are not integrated with sensory information in a linear fashion as outlined in this simple model. Regardless, the fact that the influence of action-linked expectations on perceived onset diminishes at longer delays makes it clear that more complicated mechanisms than simple statistical integration are at play during the perception of self-initiated events.

11.5 Summary

Computational models of motor control have stressed that internal models must incorporate the anticipated sensory consequences of movements in the planning of the movements themselves (Jordan & Rummelhart 1992). Our friendly waiter would not be able to do his job without being able to alter his movements based on such anticipatory processes. The emphasis in the current review is that expectations are not only important for modifying movement commands, but that these action-based signals also can influence perception. In the case of the waiter, we have shown that his perception of the weight of his tray is influenced by his own actions.

The second half of the chapter focused on one component of sensory expectations: how our actions influence the perceived time of resulting sensory events. Visual illusions such as chronostatis or the perceived duration in our vibrotactile experiments point to an influence of volitional actions on the perceived onset of sensory events initiated by those actions. Thus, one mechanism by which the action system influences perception is by influencing the perceived timing of sensory signals. The importance of timing in the generation of expectations is also evident in our hypothetical waiter example: the anticipatory response must be precisely timed. Reduce the force too early, and the postural hand will not be able to support the load. Wait until the sensory information is registered, and the load may become unstable due to the failure to reduce the supporting force.

These expectations, although adaptive, may also lead to systematic distortions of time. We introduced a simple model to capture these ideas, emphasizing how expectations and sensory information might be combined in our perception of sensory events with respect to our actions. The model provides a first pass at this integration problem, and we can already see serious limitations in its utility given the complex temporal relationship between the movements and perceived timing of the resulting stimuli. Nonetheless, it emphasizes the need to conceptualize the brain as a dynamic system in which movement- and sensory-oriented information are integrated in our perceptual experience.

Acknowledgments

We thank Joern Diedrishsen, Thom Griffiths, and Talia Konkle for helpful discussion.

References

Blakemore, S. J., Frith, C. D., & Wolpert, D. M. (1999). Spatio-temporal prediction modulates the perception of self-produced stimuli. *J Cogn Neurosci* **11**(5): 551–559.

Blakemore, S. J., Frith, C. D., & Wolpert, D. M. (2001). The cerebellum is involved in predicting the sensory consequences of action. *Neuroreport* **12**(9): 1879–1884.

Blomdell, A., Bolmsjö, G., Brogårdh, T., Cederberg, P., Isaksson, M., Johansson, R., et al. (2005). Extending an industrial root controller – implementation and applications of a fast open sensor interface. *IEEE Robotics & Automation Magazine* **12**(3): 85–94.

Charpentier, A. (1891). Analyse experimentale de quelques elements de la sensation de poids. *Archives de Physiologie Normales et Pathologiques* **3**: 122–135.

Diedrichsen, J., Verstynen, T., Hon, A., Lehman, S. L., & Ivry, R. B. (2003). Anticipatory adjustments in the unloading task: is an efference copy necessary for learning? *Exp Brain Res* **148**(2): 272–276.

Diedrichsen, J., Verstynen, T., Hon, A., Zhang, Y., & Ivry, R. B. (2007). Illusions of force perception: the role of sensori-motor predictions, visual information, and motor errors. *J Neurophysiol* **97**: 3305–3313.

Diedrichsen, J., Verstynen, T., Lehman, S. L., & Ivry, R. B. (2005). Cerebellar involvement in anticipating the consequences of self-produced actions during bimanual movements. *J Neurophysiol* **93**(2): 801–812.

Dufossae, M., Hugon, M., & Massion, J. (1985). Postural forearm changes induced by predictable in time or voluntary triggered unloading in man. *Exp Brain Res* **60**: 330–334.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415**(6870): 429–433.

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nat Neurosci* **5**(4): 382–385.

Hodinott-Hill, I., Thilo, K. V., Cowey, A., & Walsh, V. (2002). Auditory chronostasis: hanging on the telephone. *Curr Biol* **12**(20): 1779–1781.

Hugon, M., Massion, J., & Wiesendanger, M. (1982). Anticipatory postural changes induced by active unloading and comparison with passive unloading in man. *Pflugers Arch* **393**(4): 292–296.

Jordan, M., & Rumelhart, D. (1992). Forward models: supervised learning with a distal teacher. *Cog Sci* **16**: 307–354.

Kording, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature* **427**(6971): 244–247.

Lacquaniti, F., Soechting, J. F., & Terzuolo, C. A. (1982). Some factors pertinent to the organization and control of arm movements. *Brain Res* **252**(2): 394–397.

Lum, P. S., Reinkensmeyer, D. J., Lehman, S. L., Li, P. Y., & Stark, L. W. (1992). Feedforward stabilization in a bimanual unloading task. *Exp Brain Res* **89**: 172–180.

Masson, J., Ioffe, M., Schmitz, C., Viallet, F., & Gantcheva, R. (1999). Acquisition of anticipatory postural adjustments in a bimanual load-lifting task: normal and pathological aspects. *Exp Brain Res* **128**(1–2): 229–235.

Matin, L., Picoult, E., Stevens, J. K., Edwards, M. W. Jr., Young, D., & MacArthur, R. (1982). Oculoparalytic illusion: visual-field dependent spatial mislocalizations by humans partially paralyzed with curare. *Science* **216**(4542): 198–201.

Nelson, R. (1996). Interactions between motor commands and the somatic perception in sensorimotor cortex. *Curr Opin Neurobiol* **6**: 801–810.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature* **370**(6487): 256–257.

Oliver, M., Verstynen, T., & Ivry, R. B. (2003). *Did I Do That? – Modulating the Somatosensory Percept through Self Production.* Poster presented at the Cognitive Neuroscience Society meeting, 2003.

Park, J., Schlag-Rey, M., & Schlag, J. (2003). Voluntary action expands perceived duration of its sensory consequence. *Exp Brain Res* **149**(4): 527–529.

Schmidt, W. C. (2000). Endogenous attention and illusory line motion reexamined. *J Exp Psych: Hum Percept Perform* **26**: 980–996.

Sharma, J., Dragoi, V., Tenenbaum, J. B., Miller, E., & Sur, M. (2003). V1 neurons signal acquisition of an internal representation of stimulus location. *Science* **300**: 1758–1763.

Soechting, J. F., Lacquaniti, F., & Terzuolo C. A. (1986). Coordination of arm movements in three-dimensional space. Sensorimotor mapping during drawing movement. *Neuroscience* **17**: 295–311.

Toyama, K., Komatsu, Y., & Shibuki, K. (1984). Integration of retinal and motor signals of eye movements in striate cortex cells of the alert cat. *J Neurophysiol* **51**: 649–665.

van Beers, R. J., Sittig, A. C., & Denier van der Gon, J. J. (1999). Localization of a seen finger is based exclusively on proprioception and on vision of the finger. *Exp Brain Res* **125**(1): 43–49.

Weiskrantz, L., Elliott, J., & Darlington, C. (1971). Preliminary observations on tickling oneself. *Nature* **230**(5296): 598–599.

Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nat Neurosci* **5**(6): 598–604.

Welch, R. B., & Warren, D. H. (1986). Intersensory interactions. In K. R. Boff, L. Kaufman, & J. P. Thomas (eds.), *Handbook of Perception and Human Performance.* New York: Wiley.

Yarrow, K., Haggard, P., Heal, R., Brown, P., & Rothwell, J. C. (2001). Illusory perceptions of space and time preserve cross-saccadic perceptual continuity. *Nature* **414**(6861): 302–305.

Yarrow, K., Haggard, P., & Rothwell, J. C. (2004). Action, arousal, and subjective time. *Conscious Cogn* **13**(2): 373–390.

Yarrow, K., & Rothwell, J. C. (2003). Manual chronostasis: tactile perception precedes physical contact. *Curr Biol* **13**(13): 1134–1139.

12

On the perceived interdependence of space and time: evidence for spatial priming in the temporal kappa effect

GISA ASCHERSLEBEN AND JOCHEN MÜSSELER

Summary

Perceived duration of interstimulus intervals is influenced by the spatial configuration of stimuli. When participants judge the two intervals between a sequence of three stimuli presented with different spatial distances, a greater distance between two stimuli makes the corresponding time interval appear longer (kappa effect, Experiment 1). By employing a choice-reaction time task, we demonstrate that this effect is at least partly due to a facilitating influence of the preceding stimulus on the timing of the subsequent one while the timing of the first stimulus presented is not influenced by the subsequent one. Moreover, reaction times to the subsequent stimulus increased with spatial distance between successive stimuli, and this was valid for a three-stimulus condition (Experiment 2) as well as for a two-stimulus condition (Experiment 3). Thus, our results provide evidence for spatial priming in the temporal kappa effect.

12.1 Introduction

Perceiving space and time is often considered to be independent. However, the interdependency of both dimensions has been known for a long time and is most apparent in the perception of moving stimuli. For example, in 1862 Zöllner discovered a subjective spatial contraction of figures when moved behind a vertical slit (*anorthoscopic distorted pictures*, see also Vierordt 1868; Parks 1965). Through the motion, the slit uncovered only small figure sections at any time, and apparently the perceptual integration of the temporally separated sections contracted the figure spatially. This phenomenon (and related phenomena, e.g., the *Ansbacher effect*, Ansbacher 1944, or the *tandem effect*, Müsseler & Neumann 1992) demonstrates that perceived space depends on the temporal characteristics of stimulus presentation, here as a consequence of stimulus motion.

The interdependency of space and time is not only observed with moving stimuli but with stationary stimuli as well. In an early study, Benussi (1913, pp. 285) presented participants with three successive visual flashes of light at different locations, thus defining two spatial distances and two temporal intervals. Benussi found that when two distances of equal size were combined with two temporal intervals of unequal size, the distance judgments were

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

in accordance with the temporal interval. That is, shorter temporal intervals were coupled with shorter distance judgments and vice versa. Later on, this effect was also observed in the auditory (e.g., Cohen et al. 1954) and in the tactile domain (e.g., Helson & King 1931). Helson (1930) entitled this phenomenon the *tau effect*.

The tau effect stands for the influence of time on perceived space. The reverse influence, that is, the influence of space on perceived time is also documented (Cohen et al. 1954; Price-Williams 1954a). For example, in the displays of three stimuli when two intervals of equal length were combined with two distances of unequal size, the temporal judgments were in accordance with the spatial distances. That is, shorter distances were coupled with shorter time judgments and vice versa. This phenomenon was referred to as the *kappa effect* (Cohen et al. 1953). Nowadays, the tau–kappa phenomenon is a synonym for the perceived interdependence of space and time (Bill & Teft 1969, 1972; Huang & Jones 1982; Jones & Huang 1982; Sarrazin et al. 2004).

What can account for the tau–kappa phenomenon? At first sight, it seems to reflect natural laws known from modern physics. A critical insight from Einstein's Theory of Relativity was that space could not be comprehended without time and vice versa. So it was not astonishing that in the middle of the last century the psychological tau–kappa phenomenon was taken as evidence for the existence of a relativity principle that overlaps scientific disciplines (e.g., Cohen 1969). In another context, however, we have already claimed that dealing superficially with relativistic ideas is not sufficient to explain the mechanisms underlying perceptual phenomena (Müsseler 1999). For this purpose, a more strict application of the relativistic formulas would be necessary (Caelli et al. 1978; Drösler 1979). One of these formulas requires substituting c, in Einstein's physics the speed of light, with c^*, the corresponding limited velocity of transmission in the visual system. Even if c^* is operationalized and estimated in psychological terms, the empirical data of the tau-kappa experiments does not easily fit the relativistic assumptions (for details, see Müsseler 1999).

An alternative explanation that assumes the effect is based on the tendency to perceive constant velocity of apparent motion between the stimuli fits the data much better. Huang and Jones (1982), for example, found evidence that the tau effect depends on a weighted average of distance and the expected distance traversed in the given time at constant velocity (see also Jones & Huang 1982). However, Collyer (1977) and Sarrazin et al. (2004) reported data that were inconsistent with this conclusion. The authors referred to situations in which the tau–kappa phenomenon was not observed, although it should have been based on the constant-velocity hypothesis. Thus, the constant-velocity hypothesis seems to be valid only in a given/limited range of temporal and spatial stimulation.

Here we examine a different explanation of the kappa effect. It is based on a neural network account originally developed for the processing of moving stimuli (e.g., Jancke 2000; Erlhagen & Jancke 2004; see also Müsseler et al. 2002), but the underlying assumptions should be equally applicable to stationary stimuli. In consonance with Erlhagen and Jancke (2004) we assume that the presentation of a stimulus elicits a spatial activation pattern that is not restricted to the area covered by the stimulus. Rather activation spreads to and integrates contextual information from adjacent parts of the visual field. Three properties

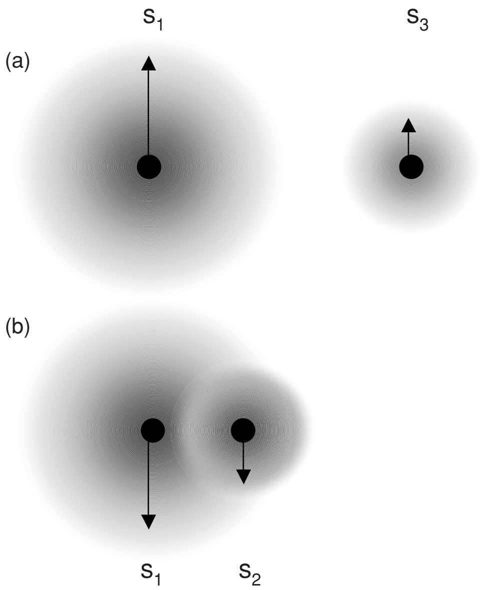


Fig. 12.1 Schematic illustration of the model assumptions. Two stimuli are presented one after the other with a large (a) and small (b) interstimulus distance. The presentation of a stimulus elicits the buildup of an activation pattern, which is not restricted to the area covered by the stimulus but rather spreads its activation to the adjacent parts of the visual field (gray gradient). The point in time illustrated when S_2 was just presented. Accordingly, the activation pattern of S_1 is more advanced than the activation pattern of S_2 (as indicated by the arrows). Both activation patterns do not affect each other with a large interstimulus distance (a), but they do with a small interstimulus distance (b). In the latter case the activation pattern of S_2 integrates and accumulates contextual information from S_1 with the consequence that the processing of S_2 speeds up.

of the account are important in the present context: (1) The buildup of the activation pattern takes time. (2) The activation pattern gradually decays with the distance from the stimulus. (3) Activation accumulates when two stimuli overlap spatially (cf. also the spatial distribution of visual attention after priming, e.g., Downing 1988; Steinman et al. 1995). This account developed for the processing of moving stimuli can also be applied to explain effects observed with spatial priming of stationary stimuli (Posner 1978, 1980).

How can these ideas account for the kappa effect? Let us consider first a two-stimulus situation illustrated in Fig. 12.1. Two stimuli are presented successively with a large (Fig. 12.1(a), i.e., S_1 and S_3) and small (Fig. 12.1(b), i.e., S_1 and S_2) interstimulus distance. The presentation of each stimulus elicits the buildup of an activation pattern, which comprises adjacent stimulus areas (gray gradient). The point in time is illustrated when S_2/S_3 was just presented. Accordingly, the activation pattern of S_1 is spatially more advanced than the one of S_2/S_3 (as indicated by the arrows). The activation patterns do not affect each other with a large interstimulus distance (a) but do with a small interstimulus distance (b). In the latter case, the activation pattern of S_2 integrates and accumulates contextual

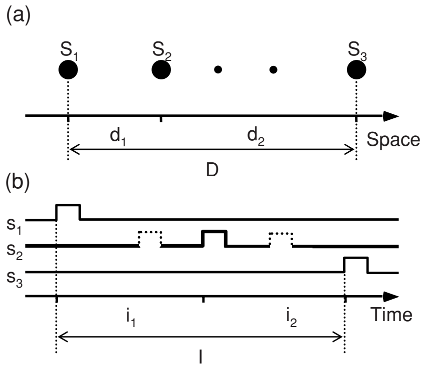


Fig. 12.2 (a) Spatial arrangement of the stimuli (S). Example of a stimulus presentation (large dots) and possible other positions (small dots). d: distance. (b) Temporal arrangement of the stimuli. i: interval.

information from S_1 with the consequence that the processing of S_2 is sped up. Thus, the spatial vicinity of the stimuli should decrease the processing time of the second stimulus (cf. the perceptual-latency-priming account, e.g., Scharlau & Neumann 2003; Scharlau 2004).

From this, it becomes immediately clear that the kappa effect should emerge from a situation with three stimuli. With temporal intervals of equal size, a reduced processing time of S_2 should result from the smaller spatial distance $S_1 - S_2$, whereas a reduction is not expected with the larger distance $S_2 - S_3$. Correspondingly, the temporal interval $S_1 - S_2$ should be perceived as being shorter than the interval $S_2 - S_3$.

The subsequent experiments aimed to examine this spatial-priming hypothesis. In Experiment 1, participants judged the temporal intervals between stimuli to establish the conditions in which the kappa effect is observed. With these conditions, participants performed choice reaction times to each of the three stimuli in Experiments 2 and 3. Choice reaction times were used as an indicator of the corresponding processing times of the stimuli.

12.2 Experiment 1

To establish the conditions in which the kappa effect can be observed, participants judged the temporal intervals between three visual stimuli. Three different spatial distances were combined with three different temporal intervals (cf. Fig. 12.2). It was expected that the temporal judgments of the intervals would depend on the interstimulus distances (kappa effect).

12.2.1 Methods

Participants. Fourteen adults participated in the experiment. In the present as well as in the subsequent experiments, all participants reported having normal or corrected-to-normal vision. Most of the participants were students at the University of Munich.

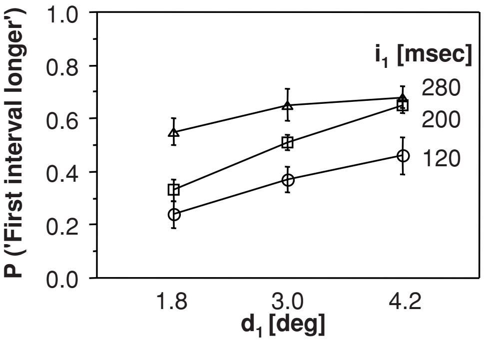


Fig. 12.3 Mean proportion (P) of first-interval-longer responses as a function of d_1 and for each level of i_1. Error bars indicate standard error between participants. Note, $d_1 + d_2$ was always 6 deg, and $i_1 + i_2$ was always 400 msec (Experiment 1).

Apparatus and stimuli. The experiment was controlled by a Macintosh computer. The stimuli were presented on a 17-inch CRT monitor with a refresh rate of 75 Hz and a luminance of approximately 40 cd/m^2. The rest of the room was dimly lit. The participants sat with their head in a forehead and chin rest facing the computer screen 50 cm in front of the monitor. We used white stimuli (dots) on a black background (0.40 deg of visual angle; presentation time was 13 msec).

12.2.2 Design and procedure

On each trial, the participant observed a sequence of three stimuli, the first and the third of which covered 6.0 deg of visual angle, that is, $D = d_1 + d_2 = 6$ deg (see Fig. 12.2). The three possible values of d_1 were 1.8 deg, 3.0 deg, and 4.2 deg. The total presentation duration was fixed at 400 msec ($I = i_1 + i_2 = 400$ msec). The three possible values of i_1 were 120 msec, 200 msec, and 280 msec (interstimulus intervals). Presentation direction was held constant from left to right.

During each block of ninety trials there were ten replications of each of the nine $d_1 - i_1$ combinations with the order of presentation being randomized. Each participant took part in six blocks. The participant responded by pressing one of two response buttons to indicate whether the first or the second temporal interval appeared to be the longer one.

The experiment began with a training block in which participants had to judge only conditions in which the spatial arrangement was symmetrical ($d_1 = d_2 = 3.0$ deg) and the temporal interval was asymmetrical ($i_1 \neq i_2$). In the training block feedback (“correct” or “wrong”) was given after each trial. Training was complete when the participants reached 85% correct answers within the last twenty trials.

12.2.3 Results and discussion

The mean proportion of responses in which the participant indicated that the first temporal interval was longer for each $d_1 - i_1$ combination is shown in Fig. 12.3. The mean values were entered into a repeated-measures analysis of variance (ANOVA) that distinguished two within-subject factors: time and distance. Both factors and the interaction reached significance [time: $F(2, 26) = 7.45$, $p = 0.003$; distance: $F(2, 26) = 22.89$, $p < 0.001$; time × distance: $F(4, 52) = 5.64$, $p = 0.001$].

The significant effect of time indicated that participants were able to clearly discriminate the different temporal intervals. Under conditions of equal spatial separation between stimuli ($d_1 = d_2 = 3.0$ deg), participants were able to distinguish the two interval durations ($i_1 = 120$ msec: $P = 0.33$ or $i_1 = 280$ msec: $P = 0.64$), whereas they were at chance ($P = 0.51$) with intervals of equal duration ($i_1 = i_2 = 200$ msec). The same was true in principle for the other two spatial conditions.

Moreover, the significant effect of the factor of distance indicated an influence of the spatial distances on the temporal judgments. When the first spatial distance was shorter than the second one ($d_1 = 1.8$ deg), participants tended to underestimate the duration of the first temporal interval, whereas they overestimated the duration of the first temporal interval when the first spatial distance was longer than the second one ($d_1 = 4.2$ deg). This effect was especially pronounced in conditions in which both intervals had the same duration ($i_1 = i_2 = 200$ msec) and somewhat less pronounced when the two intervals were not of equal length, thus leading to the significant interaction. In any case, the temporal judgments showed the expected dependency on the interstimulus distances (i.e., the kappa effect).

12.3 Experiment 2a–c

After having established the kappa effect in Experiment 1, we tested our spatial-priming hypotheses by applying a choice reaction time task. Three parallel experiments were conducted in which participants responded as fast as possible to the left or right "mouth" of a Pacman figure presented at the first stimulus position (Experiment 2a), the midposition (Experiment 2b), or the third position (Experiment 2c). The remaining two stimuli in each subexperiment were the dots presented in Experiment 1. Here, we kept the temporal intervals constant (i.e., $i_1 = i_2 = 200$ msec) and manipulated only the spatial distances by using the same values as in Experiment 1. What did we expect in each subexperiment?

In Experiment 2a, participants had to react to the Pacman figure in the first position. According to our account, the buildup of an activation pattern for a second stimulus is influenced by the activation pattern of a preceding one but not vice versa. Thus, the processing of S_1 should not be affected by the presentation of a subsequently presented stimulus.* Consequently, the manipulation of the spatial distances between S_1 and S_2 and S_2 and S_3 should not have affected choice reaction times to S_1. Alternatively, the subsequent stimulus S_2 might have masked the previous stimulus S_1, especially under close spatial conditions. Such a masking account would predict that reaction times suffer from the subsequent stimulation.

---

* There might be an influence of the second stimulus on the first one under conditions with very short temporal intervals but this is not under consideration here.

In Experiment 2b the Pacman figure was presented in the midposition (S_2). If our assumption that the previous presentation of a stimulus affects the processing speed of a subsequent one is correct, then a spatial overlap in the activation patterns should hasten processing time of S_2 under conditions with small spatial distances compared to less overlap under conditions with medium spatial distances and nonoverlap with large distances. Therefore, reaction times to S_2 were predicted to increase with an increase in spatial distance (d_1) between S_1 and S_2.

In Experiment 2c, the reverse pattern of reaction times was predicted. When the Pacman figure was in the third position, the preceding stimulus S_2 was closest when d_1 was largest. Again, in this case the buildup of an activation pattern for the stimulus (here S_3) should benefit more by the activation pattern that overlaps spatially with the previous presentation (here S_2). Thus, reaction times were predicted to increase with an increase in spatial distance (d_2) between S_2 and S_3 (and thus to decrease with an increase in spatial distance d_1).

12.3.1 Methods

Participants. Altogether forty-seven adults participated in Experiment 2, fifteen in Experiment 2a, and sixteen each in Experiments 2b and 2c.

Apparatus and stimuli. The apparatus and stimuli were the same as in Experiment 1 with the exception that the first dot (Experiment 2a), the second dot (Experiment 2b), or the third dot (Experiment 2c) was replaced with a Pacman figure (a disk with a notch) that had the notch (the "mouth") either on the right or the left side.

12.3.2 Design and procedure

The spatial features of the stimulus presentation were as in Experiment 1. However, with regard to the temporal manipulation, only the condition with identical intervals was presented ($i_1 = i_2 = 200$ msec). During each block of eighteen trials there were three replications of each of the six conditions (3 spatial conditions $\times$ 2 stimuli) with the order of presentation being randomized. Each participant took part in twenty blocks of trials.

The task of the participant was to identify whether the "mouth" of the Pacman was on the right or on the left side of the first (Experiment 2a), the second (Experiment 2b), or the third stimulus (Experiment 2c). Stimulus presentation was again always from left to right. Participants responded by pressing one of two response buttons. At the end of each block, participants received feedback about their mean reaction time and the number of errors in the preceding block. At the beginning of each subexperiment there was a training block that was not analyzed.

12.3.3 Results and discussion

Trials with reaction times lower than 100 msec and greater than 1000 msec were counted as errors. Median reaction times and percentages of errors for every participant and each

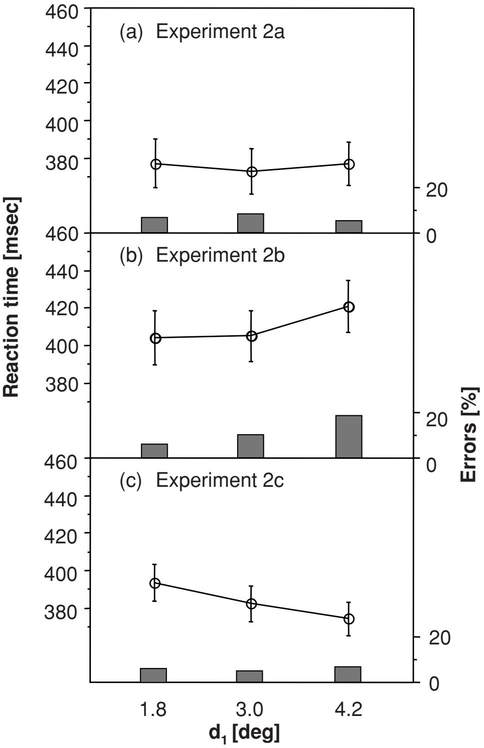


Fig. 12.4 Mean reaction times and error percentages as a function of d_1 in the choice-reaction-time task on the first stimulus (Experiment 2a), the second stimulus (Experiment 2b), and the third stimulus (Experiment 2c). In this experiment the intervals were always $i_1 = i_2 = 200$ msec.

condition were entered into two overall 3 (position of the Pacman figure) × 3 (distances d_1) ANOVAs with repeated measures on the second factor. In the ANOVA of reaction times the main effects were not significant (both $p > 0.10$). As expected, the ANOVA revealed a significant interaction ($F[4, 88] = 11.06$ and $p < 0.001$). The error analysis revealed significant main effects (position: $F(2, 44) = 4.55$, $p = 0.016$; distance d_1: $F(2, 88) = 17.03$, $p < 0.001$) and a significant interaction ($F(4, 88) = 18.97$, $p < 0.001$). To test our specific hypotheses, separate analyses were conducted for each of the three subexperiments.

In the first subexperiment (2a), the Pacman figure appeared in the left position as the first stimulus in the display. According to our assumption that the processing of the Pacman figure is only affected by previously presented stimuli, we predicted no systematic effect of distances d_1 on the reaction times.

Mean reaction times and mean percentage of errors across participants in the three different spatial conditions are shown in Fig. 12.4(a). There were only minor deviations from the grand mean reaction time of 376 msec. Accordingly, a repeated-measures ANOVA with the within-subject factor *distance* was not significant ($F[2, 28] = 1.03$, $p > 0.30$).

Mean percentages of errors were 7.0% ($d_1 = 1.8$ deg), 7.8% ($d_1 = 3.9$ deg), and 5.6% ($d_1 = 4.2$ deg). The corresponding ANOVA was significant ($F(2, 28) = 3.62$, $p = 0.040$) because of the slight reduction of errors in the condition $d_1 = 4.2$ deg. This reduction indicates that the probability for identifying the Pacman figure correctly increased slightly with the spatial separation from the other stimuli. However, the error differences between conditions were small.

In Experiment 2b, the Pacman figure appeared in the midposition as the second stimulus in the display. We predicted an increasing processing time advantage of S_2 with decreasing distance from S_1. Mean reaction times and mean error percentages are shown in Fig. 12.4(b). The ANOVA revealed a significant effect of the distances d_1 on the reaction times [$F(2, 30) = 6.97$, $p = 0.003$]. An increase in reaction times was mainly observed in the condition with the greatest d_1 (404 msec with $d_1 = 1.8$ deg, 405 msec with $d_1 = 3.0$ deg, and 421 msec with $d_1 = 4.2$ deg).

Note that the mean reaction time (410 msec) in Experiment 2b was about 35 msec higher than in Experiment 2a (376 msec). Instead, facilitation by a previously presented stimulus should lead to a reduction in overall reaction times. We suggest that the spatial uncertainty of the Pacman figure prolonged overall reaction times. Nevertheless, reaction times increased with an increase of d_1, as predicted by our hypothesis. This increase in the reaction times is accompanied by a corresponding increase in the error rates [$F(2, 30) = 30.32$, $p < 0.001$]. Errors increased significantly from 5.9% ($d_1 = 1.8$ deg) to 9.9% ($d_1 = 3.0$ deg) and 18.9% ($d_1 = 4.2$ deg). Thus, a speed–accuracy trade-off can be ruled out.

In the last subexperiment (2c), the Pacman figure appeared in the rightmost position as the third stimulus in the display. Contrary to Experiment 2b, we predicted that S_3 processing would profit most from S_2 when the distance between S_2 and S_3 was small as opposed to large.

Mean reaction times and mean error percentages are shown in Fig. 12.4(c). As expected, reaction time decreased from 393 msec ($d_1 = 1.8$ deg) to 382 msec ($d_2 = 3.0$ deg) and 374 msec ($d_1 = 4.2$ deg). Accordingly, the ANOVA showed a significant influence of the distances d_1 on the reaction times [$F(2, 30) = 16.27$, $p < 0.001$]. In this experiment, errors did not differ significantly [$F(2, 30) = 1.4$, $p > 0.30$].

In sum, the results obtained in Experiment 2a–2c support the predictions from the spatial-priming hypothesis, that is, under conditions with small interstimulus distances, the processing of a second stimulus is sped up as compared to situations with large interstimulus distances, in which no overlap of activation patterns is likely.

However, there is another factor that may have influenced the choice reaction time, at least in part: eye movements. Under the typical conditions, in which the kappa effect can be observed, eye movements are not controlled; the same holds for our Experiment 2. It might well be that participants' fixation behavior is responsible for the results, instead of spatial priming. Here it is important to remember that the spatial positions of S_1 and S_3 were always constant and, thus, predictable to the participants. To react as fast as possible to the critical stimulus, participants probably fixated its expected location. In Experiment 2a, in which the task was to react to S_1, fixation of the left stimulus resulted in S_2 and S_3 being presented in the right peripheral visual field. As predicted by the priming account, reaction time is not influenced by S_2 and S_3. But this might also be a result of the fact that stimuli presented in the peripheral visual field are easy to ignore.

On the other hand, the same strategy should have been possible in Experiment 2c, but reaction time results are different. Here, participants most probably fixated the location of S_3, the rightmost stimulus, resulting in S_1 and S_2 being presented in the left visual field. Nevertheless, the spatial distance between S_3 and the preceding stimulus S_2 had an influence on the reaction time to S_3. Thus, it is the peripheral onset and the temporal sequence of the previous events that affect processing times.

In Experiment 2b, the situation is somewhat more complicated. The spatial position of the critical stimulus S_2 varied from trial to trial. A possible strategy of the participants might have been to fixate the middle of the stimulus array. In this case, S_2 would have the same distance from the fixation point under conditions $d_1 = 1.8$ deg and $d_1 = 4.2$ deg. But with $d_1 = 1.8$ deg, S_1 and S_2 would be in the left visual field, whereas with $d_1 = 4.2$ deg, S_1 first appears in the left peripheral visual field and is then followed by S_2 appearing in the right peripheral visual field (with S_3 always appearing in the right peripheral visual field). The observed increase in reaction times to S_2 at $d_1 = 4.2$ deg might therefore be a result of S_1 and S_2 being presented in different visual fields – and not be a consequence of the spatial distance between the two stimuli. To test this alternative explanation, in the following experiment we kept the visual field, in which the stimuli were presented, constant by controlling visual fixation.

12.4 Experiment 3

The aim of Experiment 3 was to replicate the main results obtained in Experiment 2 under controlled fixation. To support the spatial-priming account, choice reaction times should decrease with decreasing spatial distance between successive stimuli; however, there should only be an influence of the preceding stimulus on the following one but not vice versa. Thus, we replicated Experiment 2b with the following important modifications. S_2 location was fixed at the center of the screen, and participants were instructed to fixate this location. As a consequence, S_1 always appeared in the left visual field and S_3 always in the right visual field with varying distances to S_2. Moreover, we tested the generalizability of the prediction that the preceding stimulus has an influence on the following one and not vice versa by comparing the three-stimulus situation with a two-stimulus condition (S_1 and S_2 only).

12.4.1 Methods

Participants. Twelve adults participated in the experiment.
Stimuli and design. The apparatus and stimuli were identical to those used in Experiment 2b, except that the second stimulus showed the Pacman figure with a notch either in the upper or lower part of the figure (requiring upper and lower key presses). Thus, the

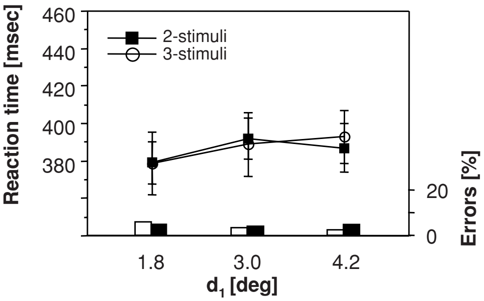


Fig. 12.5 Mean reaction times and error percentages as a function of d_1 in the choice-reaction-time task on the second stimulus (Experiment 3). Again, the intervals were always $i_1 = i_2 = 200$ msec.

left–right decision in the previous experiments was changed to an up–down one to avoid S–R compatibility effects in reaction times. Additionally, the Pacman was always presented at screen center. As a fixation mark, two vertical lines (0.4 deg each) were presented throughout the experiment 0.7 deg above and below screen center.

Two conditions were compared. In the two-stimulus condition, only S_1 and S_2 were presented. In the three-stimulus condition, all stimuli were presented as in Experiments 1 and 2. Conditions were presented blockwise with the order of blocks counterbalanced among participants.

The horizontal position of the left eye was monitored with a head-mounted and infrared light-reflecting eye-tracking device (Skalar Medical B.V., IRIS Model 6500). If a saccade was detected during the presentation of the stimulus, feedback in the form of a tone was provided. The corresponding data (1.3%) were excluded from further analyses.

12.4.2 Results and discussion

Mean reaction times and mean error percentages are shown in Fig. 12.5. A 3 (distances d_1) × 2 (two-stimuli vs. three-stimuli condition) ANOVA revealed a significant effect of the distances d_1 on the reaction times [$F(2, 22) = 7.62$; $p = 0.003$]. Collapsed across conditions, the reaction times increased from 379 msec with $d_1 = 1.8$ deg to 390 msec with $d_1 = 3.0$ deg and 4.2 deg. The main effect of condition and the interaction of it with distance were not significant ($p > .20$). A corresponding ANOVA analyzing errors revealed no significant effects.

The results again support the predictions of the priming hypothesis and rule out the alternative explanation forwarded in Experiment 2 based on an influence of the visual field on reaction times. If S_2 is presented in close spatial proximity to S_1, reaction times to S_2 are faster than if their spatial separation is large. Moreover, results revealed no difference between the two-stimuli versus three-stimuli condition, indicating a more general phenomenon. Finally, the lack of difference between conditions also supports the notion that only preceding stimuli, but not subsequent ones, have an influence on the reaction times.

12.5 General discussion

The present study investigated the effect of spatial distances between stimuli on the perceived duration of interstimulus intervals (kappa effect). In Experiment 1, the two temporal intervals between three spatially separated stimuli appeared to be longer when the spatial distance between stimuli was increased. In Experiments 2 and 3, participants performed choice reaction times to each of the three stimuli to measure the processing times of the stimuli. The results obtained in these experiments support the notion that the kappa effect originates from a facilitating influence of the preceding stimulus on the (time taken to process the) subsequent one. An influence of subsequent stimuli on the preceding ones was neither predicted nor observed.

On the basis of our choice reaction time experiments we can exclude response bias explanations of the kappa effect. Namely, participants might be influenced by the spatial context under conditions with large response uncertainty, especially if they have to judge two intervals of equal length but are forced to give a two-choice answer ("Is the first interval longer or shorter than the second interval?"). Such a response-bias explanation might explain the results obtained in time estimation tasks; however, it does not apply to the choice reaction time tasks used to measure time estimations in the present study.

The results are in line with the neural network account presented in the introduction (Jancke 2000; Müsseler et al. 2002; Erlhagen & Jancke 2004), developed for the processing of moving stimuli but now extended to stationary stimuli. The presentation of each stimulus elicits a spatial activation pattern, which is not restricted to the area covered by the stimulus but spreads its activation to and integrates contextual information from adjacent parts of the visual field. The same is true whether two (or three) stimuli are presented in relatively close spatial and temporal proximity. With a large spatial interstimulus distance, the resulting activation patterns do not affect each other. With a small interstimulus distance, the activation pattern of the second stimulus integrates and accumulates contextual information from the first stimulus with the consequence that the processing of the second stimulus is sped up. Our results show that the spatial vicinity of the stimuli decreases the processing time of the second stimulus (spatial-priming hypothesis). The postulated priming mechanism is similar to basic neurophysiological mechanisms that have been suggested to account for the processing of moving stimuli (e.g., Berry et al. 1999; Kirschfeld & Kammer 1999; Erlhagen & Jancke 2004; Jancke et al. 2004).

Converging evidence for the spatial-priming hypothesis can be found in the metacontrast paradigm. Here, two stimuli are presented at the *same* location with short interstimulus intervals. Similar to the kappa paradigm, the presentation of both stimuli should elicit spatial activation patterns, which are integrated and accumulated resulting in a speeding up of the processing of the second stimulus. Indeed, by applying a synchronization task, a predating of the mask (second stimulus) was shown but no influence of the timing of the test stimulus (first stimulus) was observed (Aschersleben & Bachmann 2007). Moreover, this effect was independent of whether the first stimulus was masked by the second one. Temporal order judgments were elicited to determine the perceived timing of the second stimulus in the metacontrast paradigm. As predicted by the priming hypothesis, various studies revealed a decrease in the processing time of the second stimulus (e.g., Bachmann 1989; Neumann et al. 1993; Aschersleben 1999; Steglich & Neumann 2000; Scharlau & Neumann 2003; Scharlau 2004). However, it is important to note that the results produced by different methods used have to be cautiously compared, at least with respect to the absolute size of the effects. Although temporal order judgments require an unsped perceptual response, reaction time tasks require participants also to program and initiate a sped motor response. Often the size of the observed effects is smaller in reaction time tasks as compared to perceptual judgment tasks, although the effects are typically in the same direction (see Aschersleben 1999, for an extended discussion).

In addition to the purely temporal interpretation of the prime's impact on the processing of the second stimulus, the priming hypothesis can be extended based on the present results. The effect of the activation pattern elicited by the first stimulus is not restricted to its spatial position, but it is determined by its position (spatial and temporal priming). With increasing spatial distance to the location of the prime, its effect declines, thus the facilitatory effect on the processing of the subsequent stimuli is reduced (cf. also Shulman et al. 1985). Moreover, as assumed by the model, the activation patterns of stimuli at different spatial positions can not only overlap in space and time, but the activation can accumulate, that is, it can result in an increase in the activation pattern of the second stimulus.

To conclude, we suggest that the kappa effect, a temporal phenomenon indicating that the perceived duration of interstimulus intervals is influenced by the spatial configuration of stimuli, results from a spatial and temporal priming of the subsequent stimuli by preceding ones via overlapping activation patterns. The suggested account is a general framework that is not limited to the kappa phenomenon but can also be applied to other phenomena observed with stationary as well as moving stimuli.

Acknowledgments

We thank Lothar Knuf for programming and Birgit Richter for her support in data collection. We also thank Beena Khurana, Marcus Baldo, and an anonymous reviewer for valuable comments on an earlier draft of the manuscript. This research was supported by a grant from the Deutsche Forschungsgemeinschaft to the two authors (DFG AS 79/3). (Correspondence concerning this manuscript should be addressed to Gisa Aschersleben, Department of Psychology, Saarland University, P.O. Box 15 11 50; D-66041 Saarbrücken, Germany. Email: aschersleben@mx.uni-saarland.de)

References

Ansbacher, H. L. (1944). Distortion in the perception of real movement. *J Exp Psychol* **34**: 1–23.

Aschersleben, G. (1999). *Aufgabenabhängige Datierung von Ereignissen* [Task-dependent Dating of Events]. Aachen: Shaker.

Aschersleben, G., & Bachmann, T. (2007). *Synchronization and Metacontrast Stimulation: Evidence for the Dual-Process Attentional Theory*. Unpublished manuscript submitted for publication.

Bachmann, T. (1989). Microgenesis as traced by the transient paired-forms paradigm. *Acta Psychologica* **70**: 3–17.

Benussi, V. (1913). Versuche zur Analyse taktil erweckter Scheinbewegungen [Experiments on the analysis of tactile apparent motions]. *Archiv für die gesamte Psychologie* **36**: 58–135.

Berry, M. J., Brivanlou, I. H., Jordan, T. A., & Meister, M. (1999). Anticipation of moving stimuli by the retina. *Nature* **398**: 334–338.

Bill, J. C., & Teft, L. W. (1969). Space-time relations: effects of time on perceived visual extent. *J Exp Psychol* **81**(1): 196–199.

Bill, J. C., & Teft, L. W. (1972). Space-time relations: the effects of variations in stimulus and interstimulus interval duration on perceived visual extent. *Acta Psychologica* **36**(5): 358–369.

Caelli, T. M., Preston, G. A., & Howell, E. R. (1978). Implications of spatial summation models for processes of contour perception: a geometric perspective. *Vision Res* **18**(6): 723–734.

Cohen, J., Hansel, C. E. M., & Sylvester, J. D. (1953). A new phenomenon in time judgment. *Nature* **172**: 901.

Cohen, J., Hansel, C. E. M., & Sylvester, J. D. (1954). Interdependence of temporal and auditory judgments. *Nature* **174**: 642–644.

Cohen, J. M. A. (1969). Relativity of psychological time. In J. M. A. Cohen (ed.), *Psychological Time in Health and Disease* (40–57). Springfield.

Collyer, C. E. (1977). Discrimination of spatial and temporal intervals defined by three light flashes: effects of spacing on temporal judgments and of timing on spatial judgments. *Perception & Psychophysics* **21**(4): 357–364.

Downing, C. J. (1988). Expectancy and visual-spatial attention: effects on perceptual quality. *J Exp Psychol Hum Percept Perform* **14**: 188–202.

Drösler, J. (1979). Relativistic effects in visual perception of real and apparent motion. *Archiv fur Psychologie* **131**(3): 249–266.

Erlhagen, W., & Jancke, D. (2004). The role of action plans and other cognitive factors in motion extrapolation: a modelling study. *Vis Cogn* **11**(2–3): 315–340.

Helson, H. (1930). The Tau Effect – an example of psychological relativity. *Science* **1847**: 536–537.

Helson, H., & King, S. M. (1931). The Tau effect: An example of psychological relativity. *J Exp Psychol* **14**: 202–217.

Huang, Y. L., & Jones, B. (1982). On the interdependence of temporal and spatial judgments. *Perception & Psychophysics* **32**(1): 7–14.

Jancke, D. (2000). Orientation formed by a spot's trajectory: a two-dimensional population approach in primary visual cortex. *J Neurosci* **20**(14): RC86.

Jancke, D., Chavane, F., Naaman, S., & Grinvald, A. (2004). Imaging cortical correlates of illusion in early visual cortex. *Nature* **428**: 423–426.

Jones, B., & Huang, Y. L. (1982). Space-time dependencies in psychophysical judgment of extent and duration: algebraic models of the Tau and Kappa effects. *Psychol Bull* **91**(1): 128–142.

Kirschfeld, K., & Kammer, T. (1999). The Fröhlich effect: a consequence of the interaction of visual focal attention and metacontrast. *Vision Res* **39**: 3702–3709.

Müsseler, J. (1999). Perceiving and measuring of spatiotemporal events. In S. Jordan (ed.), *Modeling Consciousness across the Disciplines* (95–112). Lanham, MD: University Press of America, Inc.

Müsseler, J., & Neumann, O. (1992). Apparent distance reduction with moving stimuli (Tandem Effect): evidence for an attention-shifting model. *Psychol Res* **54**: 246–266.

Müsseler, J., Stork, S., & Kerzel, D. (2002). Comparing mislocalizations with moving stimuli: the Fröhlich effect, the flash-lag effect, and representational momentum. *Vis Cogn* **9**: 120–138.

Neumann, O., Esselmann, U., & Klotz, W. (1993). Different effects of visual-spatial attention on response latency and temporal-order judgment. *Psychological Research/Psychologische Forschung* **56**: 26–34.

Parks, T. E. (1965). Post-retinal visual storage. *Am J Psychol* **78**: 145–147.

Posner, M. I. (1978). *Chronometric Explorations of Mind.* Oxford, England: Erlbaum.

Posner, M. I. (1980). Orienting of attention. *Q J Exp Psychol* **32**: 3–25.

Price-Williams, D. R. (1954a). A further study of space-time perception. *Tohoku Journal of Experimental Psychology* **1**: 39–44.

Price-Williams, D. R. (1954b). The kappa effect. *Nature* **4399**: 363–364.

Sarrazin, J. C., Giraudo, M. D., Pailhous, J., & Bootsma, R. J. (2004). Dynamics of balancing space and time in memory: Tau and Kappa effects revisited. *J Exp Psychol: Hum Percept Perform* **30**(3): 411–430.

Scharlau, I. (2004). The spatial distribution of attention in perceptual latency priming. *Q J Exp Psychol: Hum Exp Psychol* **57A**: 1411–1436.

Scharlau, I., & Neumann, O. (2003). Perceptual latency priming by masked and unmasked stimuli: evidence for an attentional interpretation. *Psychological Research/Psychologische Forschung* **67**: 184–196.

Shulman, G. L., Wilson, J., & Sheehy, J. B. (1985). Spatial determinants of the distribution of attention. *Perception & Psychophysics* **37**: 59–65.

Steglich, C., & Neumann, O. (2000). Temporal, but not spatial, context modulates a masked prime's effect on temporal order judgement, but not on response latency. *Psychol Res* **63**: 36–47.

Steinman, B. A., Steinman, S. B., & Lehmkuhle, S. (1995). Visual attention mechanisms show a center-surround organization. *Vision Res* **35**: 1859–1869.

Vierordt, K. (1868). *Der Zeitsinn nach Versuchen [The sense of time in experiments].* Tübingen (Germany): Laupp.

Zöllner, F. (1862). Ueber eine neue Art anorthoskopischer Zerrbilder [About a new kind of anorthoscopic distorted figures]. *Annalen der Physik. Poggendorfs Annalen* **117**: 477–484.

Part III

Temporal phenomena: binding and asynchrony

13

Dynamics of visual feature binding

COLIN W. G. CLIFFORD

Summary

This chapter is concerned with the temporal aspects of visual binding. In particular, it concentrates on findings from studies of perceptual asynchrony between stimulus features and the temporal resolution of feature binding. I review the circumstances in which perceptual asynchronies are apparent versus those in which they are not. I argue that the existing data cannot be accounted for simply by a characteristic latency difference in the processing of different visual attributes (Moutoussis & Zeki 1997a,b) or by a scheme of temporal markers at salient stimulus transitions (Nishida & Johnston 2002). Instead, I outline a potential mechanism based on feedback from higher visual areas to primary visual cortex to account for the dynamics of binding color with orientation and direction of motion.

13.1 Introduction

How is the content of our conscious visual experience related to neural processing? Is our visual awareness an online monitor of visual processing, or do interpretative processes intervene to give conscious visual experience a postdictive quality? In the words of William James, "A succession of feelings, in and of itself, is not a feeling of succession. And because, to our successive feelings, a feeling of their own succession is added, that must be treated as an additional fact requiring its own special elucidation" (James 1890). But what is the nature of this "additional fact"? The simplest account would seem to be that the perceived sequence of events is directly related to the amount and duration of neural processing needed to achieve conscious experience (Jeannerod 1992). However, according to Dennett and Kinsbourne (1992), the experienced sequence of events is not simply a direct reflection of perceptual latencies but rather a product of the brain's "interpretative processes."

Although issues such as these have traditionally been the province of philosophers, recent visual psychophysical studies have opened them up to empirical investigation. Here I present a selective review of these studies in the context of visual feature binding and propose an account of the relationship between the dynamics of our subjective visual awareness and the time course of the underlying neural processes.

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

13.2 Modularity and binding

Substantial evidence indicates that the processing of visual information in extrastriate cortex is to some extent modular in character (see Zeki 1993). For example, primate area MT/V5 contains a high proportion of direction-selective cells that show only weak chromatic input (Seidemann et al. 1999) as one might expect from a motion-processing module. According to Bartels and Zeki (1998), Zeki and Bartels (1998, 1999), and Zeki (2003), activity within such cortical processing modules generates a "micro-consciousness" for the attribute for which that module is specialized, such that there are a number of functionally specialized microconsciousnesses corresponding to the activity of cells within different processing systems. So area MT/V5, for example, would be considered not just the motion-*processing* center in primate visual cortex but also the motion *perception* center. In this way, the neural activity underlying conscious visual perception would be spatially distributed across the various processing structures within the visual cortex.

If these microconsciousnesses operate essentially autonomously, as Zeki and Bartels (1998) propose, then how might they remain in temporal registration so that, at any given moment, activity across the visual cortex refers to the same temporal snapshot of the visual world? Here, Zeki and Bartels (1998) make a controversial proposal – perhaps there is no mechanism of temporal registration. Perhaps, instead, each microconsciousness operates with its own characteristic latency such that, at any given instant, the contents of the different microconsciousnesses represent temporally distinct snapshots of the visual world. If this were the case, visual consciousness would be distributed across the cortex not only spatially but also temporally, with each microconsciousness operating asynchronously from the others.

Modularity brings with it a binding problem. If different attributes of a scene are processed by relatively autonomous processing modules, then how is the information provided by these modules "bound" together to give rise to the experiential unity of visual consciousness? The Zeki and Bartels (1998) proposal suggests that the dynamics of binding the neural activity generating different microconsciousnesses should depend at least in part upon the dynamics of processing in the modules whose activity is being bound. If, for example, color and motion were processed in different modules, each with a characteristic processing time, then the perceptual binding of color and motion should show an asynchrony characteristic of the differential processing latency between the two attributes (Moutoussis & Zeki 1997b).

13.3 A psychophysical paradigm to study the dynamics of visual binding

The idea of a temporally distributed consciousness was motivated in large part by psychophysical findings from Moutoussis and Zeki (1997a,b) that appear to demonstrate an asynchrony between the perception of color, motion, and orientation. In their experimental paradigm, two stimulus attributes, such as motion and color, vary periodically over time between two states (e.g., up and down for motion; red and green for color). The rate of

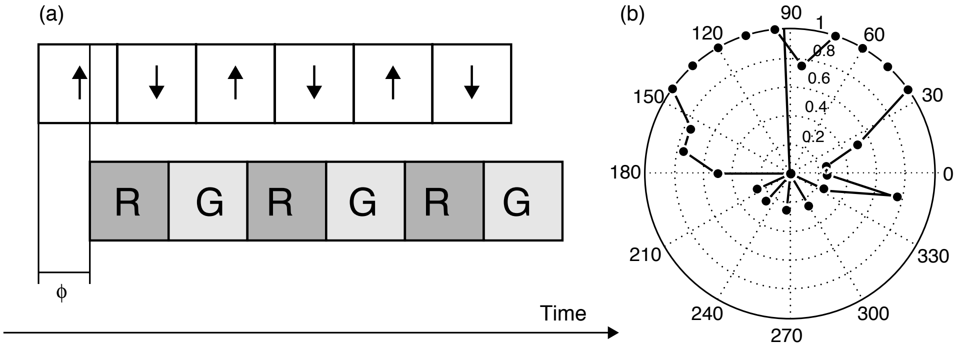


Fig. 13.1 (a) Schematic of stimulus sequence used to study visual binding. A single stimulus alternates between two colors (e.g., red [R] and green [G]) and two directions of motion (e.g., up and down). The alternations in color and motion have the same period, but their relative phase, φ, can vary from trial to trial. Subjects make a forced-choice judgment as to which color is paired predominantly with which direction of motion. (b) Example data for visual binding of color and motion (Clifford, Spehar, & Pearson 2004). Data points represent the proportion of trials a given color and direction (see scale to right of vertical radial axis) were paired together at each of thirty values of φ. The centroid representing perceptual synchrony is close to a phase of 90 deg (corresponding to 125 msec for the oscillation period of 500 msec), even though at this phase there is no physical correlation between color and direction of motion.

oscillation of each stimulus attribute is the same (e.g., 1.0 Hz), but the relative phase (φ) of the oscillations varies from trial to trial (Fig. 13.1(a)).

When $\varphi = 0$ deg, color changes in phase with motion so that upward motion is physically paired with red and downward with green. In this case the correlation (ρ) between upward motion and the color red is 1. When $\varphi = 180$ deg, the changes are still physically synchronous but the pairing is reversed ($\rho = -1$). When $\varphi = 90$ deg or 270 deg the changes between color and motion are perfectly out of phase ($\rho = 0$). In a forced-choice task, subjects are required to report which of the two motion states coexist with a particular color state, for example, which way red is going. To establish if there is an asynchrony between the perception of color and the perception of motion, the proportion of times that red is reported as going upward is recorded as a function of the relative phase, φ. If the temporal binding of color and motion is veridical, then the distribution of reported coexistence of upward motion and red color will be centered on $\varphi = 0$ deg. If an asynchrony exists between the perception of one attribute relative to the other, the distribution will be shifted in phase (Fig. 13.1(b)).

Moutoussis and Zeki (1997a) found that subjects' responses were centered on a phase difference of 39 deg for an oscillation period of 716 msec, showing that, on average, color was perceived ahead of motion by 78 msec. They interpreted the apparent latency in perception of motion relative to color as evidence of a latency difference in the processing of motion relative to color. However, the interpretation of these results in terms of a differential latency in processing is controversial for several reasons. First, a processing lag of motion behind color appears to contradict neurobiological evidence suggesting that, if anything, motion should be processed faster than color (Schiller & Malpeli 1978; Munk et al. 1995). Second, the existence of interattribute perceptual asynchronies depends on task and stimulus parameters (Arnold & Clifford 2002; Bedell et al. 2003; Clifford et al. 2003; Clifford, Spehar et al. 2004). For example, reaction times to changes in color and motion have been reported not to differ (Barbur et al. 1998; Nishida & Johnston 2002). Third, processing time may not correlate directly and invariantly with perceived time of occurrence (Johnston & Nishida 2001). Instead, it is possible that simultaneous neural activity might be consciously experienced as being representative of events that have occurred at different times (Dennett & Kinsbourne 1992), as has been proposed for the flash-lag illusion (Eagleman & Sejnowski 2000).

13.4 Using adaptation to probe the dynamics of the processes underlying binding

The use of subjective reports to examine asynchronies in the perception of different visual attributes cannot reveal whether the root of the asynchronies lies in a differential latency for processing, say, motion relative to color, or whether it is a result of subsequent interpretative mechanisms. Arnold et al. (2001) examined this issue by making use of the "psychologist's microelectrode" adaptation (Frisby 1980). Adaptation has traditionally been used to investigate the mechanisms of sensory coding (e.g., Clifford 2005), although recently its utility as a technique to investigate the neural correlates of visual consciousness has become widely recognized (He et al. 1996; Humphrey & Goodale 1998; He & MacLeod 2001; Rajimehr 2004; Shady et al. 2004; Blake & He 2005).

To probe the dynamics of color and motion processing, Arnold et al. (2001) used as a tool the motion aftereffect (MAE; Mather et al. 1998). The MAE is a result of prolonged viewing of a pattern in constant motion. The MAE is such that a subsequently viewed static pattern appears to drift in the direction opposite to the motion of the adapting pattern. The direction of the MAE can be made contingent upon the color of the moving stimulus giving rise to the color-contingent MAE (Favreau et al. 1972). For example, after prolonged exposure to red and green patterns moving in opposite directions, a stationary red pattern will appear to move in the direction opposite to the adapting motion of the red pattern while a stationary green pattern will appear to move in the direction opposite to the adapting motion of the green pattern.

The adapting and test stimulus used by Arnold et al. (2001) was a radial grating presented in a circular aperture. In the adaptation phase, the duty cycle was similar to that used by Moutoussis and Zeki (1997a) in that both the direction of motion and color reversed periodically (at 1.0 Hz). Motion alternated between clockwise and counterclockwise rotation, whereas the color of the grating changed from red–black to green–black. By correlating color states disproportionately with two directions of motion in the adapting stimulus, color-contingent MAEs were produced and measured as a function of the range of physical correlations (relative phases were sampled in steps of 36 deg). The initial adaptation period

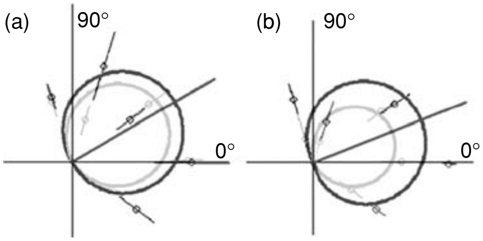


Fig. 13.2 Signed magnitude of color-contingent motion aftereffect as a function of the phasic relationship in the adapting stimulus for (a) subject DA (b) subject CC. Data are replotted from Arnold et al. 2001. Aftereffect magnitude is represented as distance from the origin. Aftereffect direction is indicated by the color of the data point (e.g., dark/light gray data point on horizontal axis represents aftereffect to an adapting phase of 0/180°). Data have been fitted with a sinusoid using three free parameters (phase, amplitude, and a constant bias term). Dark and light gray portions of the curve show the fit to the dark and dark and light gray data points, respectively. For both subjects, the best fitting sinusoid is shifted in phase (oblique line) from physical synchrony (phase of 0°) in a manner consistent with a processing advantage for color over motion.

was 5 min, with a further 4 sec of top-up adaptation after each trial. Only one phasic relationship of the adapting stimulus was used within any one run, and each top-up adaptation period commenced at a random point within the stimulus sequence to ensure that any aftereffects could not be a result of a systematic point of offset of the adapting stimulus.

In the test phase, subjects were required to indicate the direction of motion of a red–black or green–black grating rotating slowly either clockwise or anticlockwise. Psychometric functions were then fitted separately to the data for the two different color gratings to give measures of the speed of subjective stationarity for each. The strength of the contingent aftereffect was taken to be the difference between the signed speeds of subjective stationarity for the two colors minus any difference evident from an unadapted baseline condition.

Measuring aftereffect strength as a function of relative stimulus phase, φ, allowed estimation of the processing asynchrony between color and motion from the centroid of the distribution of contingent aftereffects. If color and motion were processed synchronously, the distribution should be centered on $\varphi = 0$ deg. The centroids of the data from the two subjects were at 21 deg and 33 deg, with color leading motion (Fig. 13.2). As the period of one cycle (360 deg of phase) was 1 sec, this phase difference corresponds to processing advantages for color of approximately 75 msec, comparable to the perceptual asynchronies typically observed between color and motion.

The finding that implicit processing manifests a similar asynchrony to conscious report argues against the theory that apparent perceptual asynchronies could be due entirely to attentional shifting (Enns & Oriet 2004) or other cognitive factors. Instead, it is consistent with the existence of a direct relationship between the perceived timing of events and the dynamics of the underlying neural processing.

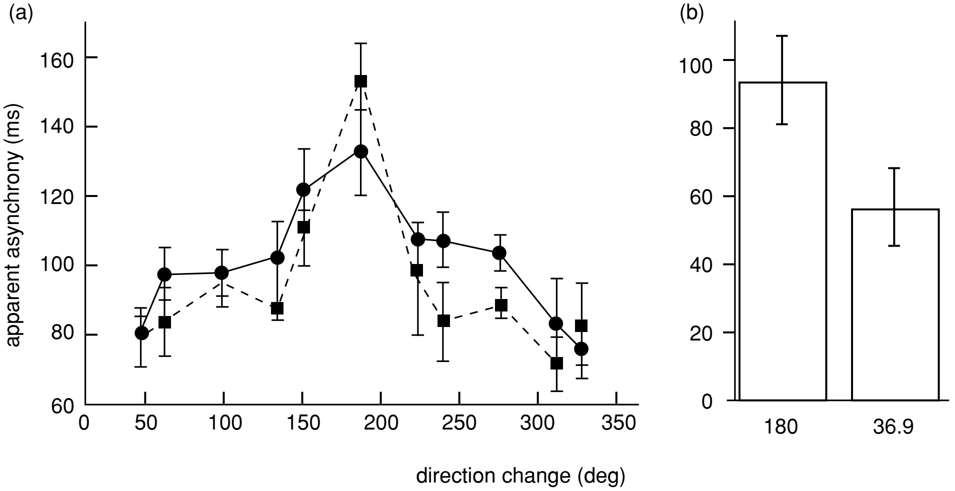


Fig. 13.3 Perceptual asynchrony between color and motion as a function of the magnitude of the change in direction between alternating motions. Data from (a) subjects DA (circles) and CC (squares) and (b) eight subjects naïve to the purposes of the study. Asynchrony was shown to vary systematically as a function of the magnitude of the direction change with the largest asynchrony occurring for changes of 180 deg (Arnold & Clifford 2002).

13.5 Determinants of perceptual asynchrony between color and motion: angle of direction change

Further evidence for a close link between the time course of perceptual processing and explicit perception comes from studies by Arnold and Clifford (2002), Bedell et al. (2003), and Clifford, Spehar et al. (2004). As in the Moutoussis and Zeki (1997a,b) studies, color was paired with motion. However, instead of always alternating between motion in opposite directions, the angular difference between the two motions was varied. These studies found the perceptual asynchrony between color and motion to be greatest for alternation between opposite motions (Fig. 13.3). The psychophysical studies suggest that perceptual asynchronies between color and motion reflect a difference in the dynamics of the underlying processing difference due in part to opponent motion inhibition.

The dependence of perceptual asynchrony on the angular difference between alternating directions of motion poses a problem for temporal marker accounts (see Nishida & Johnston, this volume) because it is not clear why the position of a temporal marker signaling a given direction of motion should depend on the magnitude of the preceding direction change. Instead, it is consistent with a delayed response to motion following direction change due to direction-selective inhibitory mechanisms (Arnold & Clifford 2002; Bedell et al. 2003; Clifford, Spehar et al. 2004). According to this account, the perceptual lag of motion relative to color is greater for 180 deg than lesser direction changes because of opponent motion inhibition. Although directional tuning of inhibition has been measured in macaque area

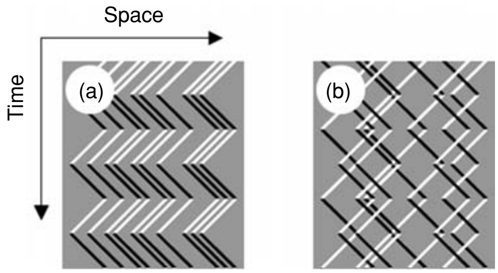


Fig. 13.4 Space–time plots of stimuli alternating in contrast polarity and direction of motion used by Clifford, Spehar, and Pearson (2004). (a) Dots oscillate from the same initial phase within a given duty cycle, all undergoing the same reversal in direction of motion at the same time. (b) Half of the dots start at one point in the duty cycle, with the rest starting at the opposite phase. All the dots change in a given attribute at the same time (e.g., for motion, half change from left to right while the other half change from right to left). It is interesting to note that in the space–time plots, as in the moving stimuli themselves, there is a much stronger tendency for the trajectories of a given contrast and orientation in space–time (velocity) to group together perceptually for the stimulus in (b), overriding the perception of individual dot trajectories alternating in direction of motion. In the moving stimulus, this grouping corresponds to the perception of motion transparency.

MT for simultaneously presented motions (Snowden et al. 1991), the physiological basis of any such effect in the temporal domain remains uncertain.

13.6 Determinants of perceptual asynchrony between color and motion: motion transparency

Motion transparency is observed when, for example, patterns of dots moving in opposite directions are superimposed. The percept is one of two sheets of dots sliding transparently across one another (van Doorn & Koenderink 1982). Several recent studies have investigated the relationship between the perception of motion transparency and the binding of color and motion (Clifford, Spehar et al. 2004; Moradi & Shimojo 2004; Wu et al. 2004; Arnold 2005). The basic stimulus used by Clifford, Spehar et al. (2004) was a field of random dots that alternated between leftward and rightward motion with a period of 500 msec. The dots also alternated in contrast polarity or color relative to the uniform gray background with the same period. The relative phase of the contrast changes and the motion changes was varied to allow measurement of the perceptual asynchrony, just as in the Moutoussis and Zeki (1997a,b) studies described earlier.

In all previous studies, dots oscillated from the same initial phase within a given duty cycle, all undergoing the same reversal in direction of motion at the same time (Fig. 13.4(a)). Clifford, Spehar et al. (2004) introduced a condition in which half of the dots started at one point in the duty cycle, with the rest starting at the opposite phase (Fig. 13.4(b)). Here the perceptual asynchrony nearly vanished, amounting to 20 msec or less. In the latter condition, both contrast polarities and both directions of motion are present at all times, and the percept is one of motion transparency (Kanai et al. 2004) such that dots moving in

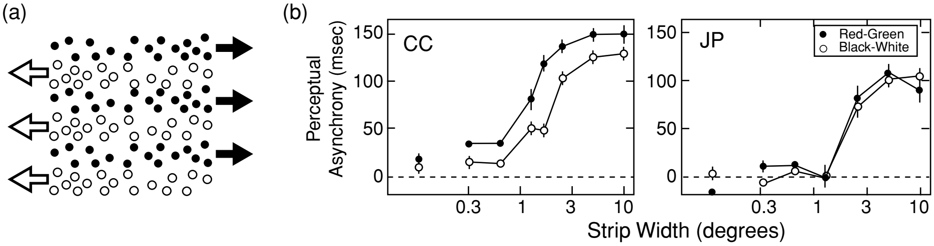


Fig. 13.5 The effect on perceptual color–motion asynchrony of dividing the stimulus into strips (Clifford, Spehar, & Pearson 2004). (a) Schematic of stimulus divided into strips parallel to the axis of motion. Dots in alternate strips started at opposite phases in the same duty cycle. Thus, on any given frame, dots in alternate strips differed in both color and direction of motion. (b) Measured perceptual asynchrony for two observers as a function of strip width for dots changing color between red and green or black and white. The stimulus subtended 10 × 10 deg of visual angle, so a strip width of 10 deg corresponded to a homogeneous stimulus. Symbols at the far left of each graph represent the results from trials where dots starting at opposite phases in the duty cycle were spatially interleaved rather than divided into strips.

a given direction are perceptually grouped together, overriding the perception of individual dot trajectories alternating in direction of motion.

As well as varying the distribution of initial phases across dots, Clifford, Spehar et al. (2004) investigated other ways to manipulate the perception of transparency in oscillating dot displays. In one experiment, all dots again alternated between leftward and rightward motion and between light and dark, with one half having one phase and the other half having the opposite phase. However, instead of their spatial arrangement being entirely random as in the previous experiment, the two groups of dots were arranged in alternating horizontal strips (Fig. 13.5). At large strip widths, the different phase relationships caused the strips to segment perceptually. When the strips were narrow, dots in nonadjacent strips but moving in the same direction grouped together, and the stimulus was perceived as two transparent surfaces moving in opposite directions (van Doorn & Koenderink 1982). For wide strips, perceptual asynchrony was considerable (around 100 msec). However, the asynchrony decreased as strip width was reduced and was again nearly abolished for strip widths less than 2 deg (about the width of a foveal MT receptive field; Albright & Desimone 1987). Manipulations of stereoscopic disparity and relative speed were also used to create and abolish transparency, and again perceptual asynchrony greatly diminished whenever transparency occurred.

The phenomenology of color–motion binding presents us with an apparent paradox. When all dots oscillate in phase (Fig. 13.4(a)), binding color and motion might seem trivial. For example, one might imagine the visual system taking an instantaneous “snapshot” of the stimulus attributes to recover the actual pairing of color and motion. However, in this situation Clifford, Spehar et al. (2004) found that perceptual asynchrony is both largest and most variable (within subjects), indicating that the problem of binding color and motion is far from straightforward. Conversely, when the dots oscillate at different phases (Fig. 13.4(b)), the binding problem might appear to be at its hardest. Any snapshot of stimulus attributes would contain both colors and both directions of motion. But in this case there is little or no perceptual asynchrony, and the task seems almost effortless. How, then, might the binding problem between color and motion be solved in transparent motion?

13.7 A role for feedback in binding color and motion

The stimuli used by Clifford, Spehar, et al. (2004) were constructed to be identical at the level of individual dots, so spatial pooling of similar motion signals would be required to differentiate coherent from transparent motion. Although neurons in V1 show little or no difference in their response to coherent versus transparent motion, marked response differences are evident in MT (Snowden et al. 1991; Qian & Andersen 1994) where receptive fields are larger (Albright & Desimone 1987). Indeed, it is probably because both directions of motion occur within a single MT receptive field, but not within a single V1 receptive field, that the stimulus appears transparent (Qian & Andersen 1994; Qian et al. 1994a,b). Consequently, the spatial resolution necessary for accurate binding of color and motion for individual dots in transparent motion has been lost by the time visual information reaches MT. Thus, if the mechanisms of motion transparency facilitate perceptual binding, the mechanism of this facilitation would appear to involve the feedback of information from MT to cortical areas with higher spatial resolution such as V1. Such feedback connections have been shown to exist anatomically (Shipp & Zeki 1989) and have been proposed to be involved in perceptual binding (Hochstein & Ahissar 2002) and in gating the contents of visual awareness (Pascual-Leone & Walsh 2001).

Feedback is envisaged to play a role in binding color and motion as follows (Fig. 13.6(a)). Signals from early cortical neurons selective for local image motion are transmitted to modular extrastriate areas where they are grouped into sustained surface representations on the basis of the common motion they represent. These areas feed back signals to earlier neurons that share their selectivity for motion. In this way neurons in early cortex are reciprocally associated with the maintained extrastriate representation of the motion-defined surfaces to which they contribute. This reciprocal association serves to bind the chromatic tuning of neurons in early cortex to the motion-defined surfaces to which they contribute. For example, attentional selection of a particular surface in turn selects the neurons in early cortex associated with that surface. These neurons feed into higher color-processing areas whose activity determines the perceived surface color. This account is somewhat speculative. For example, physiological evidence suggests that V1 neurons selective for both color and motion are comparatively rare (Horwitz & Albright 2005). However, it does at least offer a mechanism through which the binding problem between color and motion might be solved for transparent motion.

How would such an account predict the existence of marked perceptual asynchronies between color and motion when all dots oscillate in phase? A stimulus with all dots oscillating from the same initial phase (Fig. 13.4(a)) would be expected to give rise to phasic

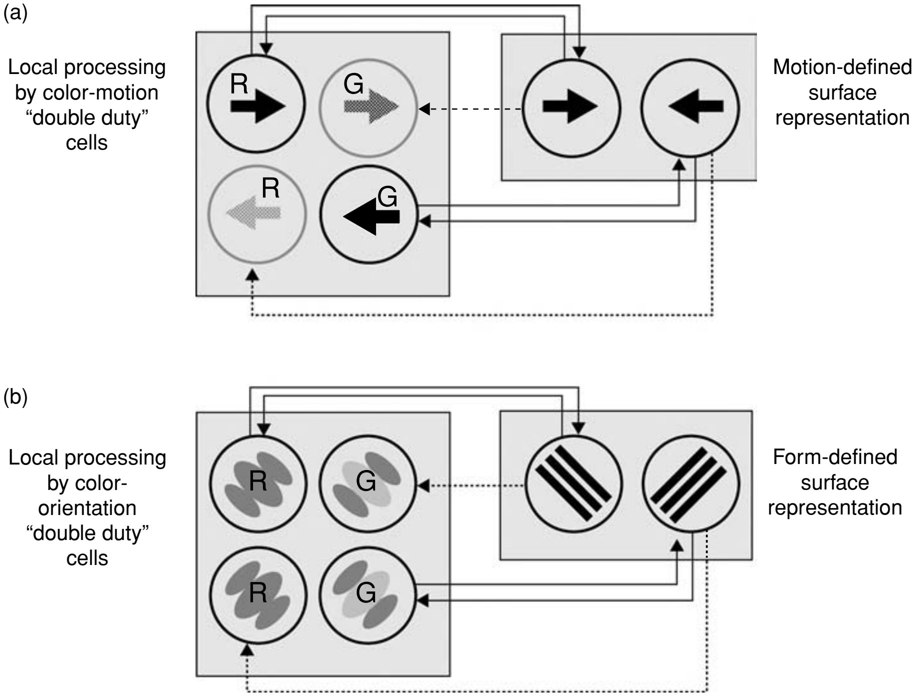


Fig. 13.6 Schematic illustrating the proposed role of extrastriate feedback in the binding of (a) color and motion during motion transparency and (b) color and orientation during temporal transparency. Signals from early cortical neurons selective for local image motion/orientation are transmitted to modular extrastriate areas where they are grouped into sustained surface representations. These areas feed back signals to earlier neurons, which share their selectivity for motion/orientation. In this way the chromatic tuning of the responsive neurons in early cortex is bound to the motion-/form-defined surfaces, determining their perceived color.

responses in populations of motion-selective neurons as found in area MT, corresponding to the dynamic formation and dissolution of motion-defined surface representations. These phasic responses would be subject to directional inhibition such that the response latency to a new direction of motion would be a function of the magnitude of the preceding direction change, delaying the formation of each new surface representation. Feedback from motion-selective areas to early visual cortex would then be similarly delayed relative to chromatic signals due to directional inhibition, giving rise to an asynchrony between the perception of the motion-defined surface and the color to which it is bound. However, when not all dots oscillate from the same initial phase (Fig. 13.4(b)), the responses of motion-selective populations should be more or less constant (Snowden et al. 1991; Qian & Andersen 1994), corresponding to the maintained representation of two transparent motion-defined surfaces. The feedback of a constant signal should not bias the temporal binding of color and motion, which would thus be close to veridicality.

A role for feedback to V1 is also consistent with the asynchrony evident in the color-contingent MAE (Arnold et al. 2001). Although the neural loci of the mechanisms mediating the various forms of contingent adaptation are not certain, the color-contingent MAE shows little or no interocular transfer (Favreau et al. 1972; Mayhew & Anstis 1972) and can be elicited by adaptation to a locally paired opposite-motion dot display (Blaser et al. 2005), suggesting that it probably reflects changes in the activity of cells in V1. If extrastriate areas use feedback to V1 to "gate" their inputs, then the timing of this gating might be important for adaptation of "double-duty" units in V1 selective for both the color and direction of motion of the stimulus.

13.8 The dynamics of binding color and orientation

Early studies concerning the perception of stimuli alternating in color and orientation yielded apparently contradictory results, prompting conflicting interpretations. Using the psychophysical paradigm described previously, Moutoussis and Zeki (1997b) found a perceptual asynchrony between color and orientation of around 63 msec, with color leading. This was taken as evidence that different processing systems for color and orientation create their corresponding percepts independently and with different delays (Moutoussis & Zeki 1997b; Zeki & Bartels 1998). However, it was subsequently reported that color and orientation can be correctly paired at oscillation rates of up to 18.8 Hz (Holcombe & Cavanagh 2001), corresponding to a period of 53 msec. Given that a processing delay of one-quarter of a temporal cycle or more would lead to systematic errors in pairing, this would seem to place an upper bound of around 13 msec on the time by which the processing of color could lead that of orientation. Instead, Holcombe and Cavanagh (2001) proposed that color and orientation are coded in combination explicitly by early stages of the visual hierarchy. The initial empirical evidence therefore seemed to provide a paradox between an apparent temporal advantage for color relative to orientation (Moutoussis & Zeki 1997b) and the ability to pair in-phase oscillations in color and orientation at high rates of alternation (Holcombe & Cavanagh 2001).

Clifford et al. (2003) investigated the dynamics of binding color and orientation using the method of Moutoussis and Zeki (1997a,b) but systematically varying the period of stimulus alternations (Fig. 13.7). For the shortest period examined, 100 msec, the measured perceptual asynchrony for all three observers was less than 10 msec, with any asynchrony tending to be a lead of orientation over color. However, at the longest period (1 sec) the perceptual asynchrony was around 50 msec, with color leading orientation. Thus, at slow alternation rates, Clifford et al. (2003) found a perceptual asynchrony of the same sign (color leading orientation) and similar magnitude to that reported by Moutoussis and Zeki (1997b). But this asynchrony was not evident at faster alternation rates, consistent with the finding of Holcombe and Cavanagh (2001) that correct pairing of alternating colors and orientations is possible at rates of nearly 20 Hz.

How can we reconcile these findings? If the perceptual asynchrony were due to a difference in the characteristic processing latencies of color and orientation (Moutoussis & Zeki

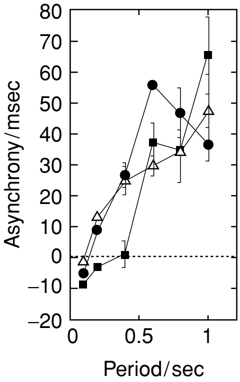


Fig. 13.7 Measurements of the perceptual asynchrony between color and orientation for three subjects as a function of oscillation period (Clifford et al. 2003). Positive values indicate a perceptual lead of color over orientation.

1997b), then it should also be evident at high alternation rates, which is clearly not the case (Holcombe & Cavanagh 2001; Clifford et al. 2003). On the other hand, if correct pairing of color and orientation at high alternation rates is a consequence of their being coded in combination early in the visual hierarchy (Holcombe & Cavanagh 2001), then how are we to account for the existence of considerable perceptual asynchronies at slower alternation rates (Moutoussis & Zeki 1997b; Clifford et al. 2003)?

13.9 A role for feedback in binding color and orientation

Physiological and psychophysical evidence suggest that color and orientation are coded in combination at the earliest levels of cortical processing (Johnson et al. 2001; Forte & Clifford 2005). However, binding of color and orientation at this early level need not necessarily rule out a contribution of feedback from higher visual areas. Propagation of signals from one area to another takes a finite amount of time and is likely to add temporal variability (although see Shadlen & Newsome 1998). Consequently, the temporal profile of the signals reaching and returning from higher visual areas might be expected to resemble delayed and dispersed (low-pass filtered) versions of the temporal stimulus waveform. I propose that interattribute binding is supported by early cortical mechanisms selective for both color and orientation but that the dynamics of this binding are influenced by feedback from higher visual areas (Fig. 13.6(b)) in a similar way to the binding of color and motion (Fig. 13.6(a)).

A stimulus alternating in color and orientation would be expected to give rise to phasic responses in populations of orientation-selective neurons along the visual hierarchy. At the higher levels of the visual hierarchy, this phasic activity might correspond to the dynamic formation and dissolution of orientation-defined surface representations. Propagation of orientation information up to higher visual areas would delay the formation of each new surface representation. Feedback from these higher areas to early visual cortex would then introduce a further delay relative to the chromatic signals, giving rise to an asynchrony between the perception of the orientation-defined surface and the color to which it is bound. However, at high alternation frequencies (faster than about 8 Hz), the percept of alternating oriented patterns gives way to one of temporal transparency (Holcombe 2001; Holcombe & Cavanagh 2001), whereby the two gratings are experienced as coexisting in space and time. Similar to motion transparency (Snowden et al. 1991; Qian & Andersen 1994), this temporal transparency may be the result of maintained representations of two orientation-defined surfaces at higher levels of the visual system. If the responses of these higher-level orientation-selective populations are indeed more or less constant, then feedback from this level should not bias the temporal binding of color and orientation, which would thus be close to veridicality as is the case empirically at alternation rates of around 10 Hz (Clifford et al. 2003).

In the case of a visual attribute, such as global form, that can be extracted with high temporal resolution (Clifford, Holcombe et al. 2004) but is not represented in early visual cortex, this account correctly predicts that high temporal resolution binding with color is not possible. Instead, binding of the global form in alternating Glass patterns with the color of the constituent dots is limited to rates no greater than 4–5 Hz (Clifford, Holcombe et al. 2004). Such low temporal resolution binding is presumably mediated by alternative mechanisms not reliant upon early conjunctive coding of attributes.

13.10 A general role for feedback in visual feature binding?

The previous sections have proposed a role for the feedback of information from extrastriate cortex to V1 in the binding of motion and orientation with color. In each case, color has been assumed to take on a secondary role, with the magnitude of the perceptual asynchrony essentially determined by the dynamics of processing and transmitting information about motion- or orientation-defined surfaces. However, a systematic perceptual asynchrony is also evident in the binding of motion and orientation (Moutoussis & Zeki 1997b; Viviani & Aymoz 2001). Indeed, the magnitude of the motion–orientation perceptual asynchrony has been reported to be approximately equal to the difference in perceptual asynchronies between motion color and orientation color (Moutoussis & Zeki 1997b). This suggests that the magnitude of the perceptual asynchrony is related to some aspect of the relative speed of processing of the attributes being bound. In the context of the proposed role of feedback in visual feature binding, it would seem arbitrary to assume that orientation would take on a primary role when bound with color but a secondary role when bound with motion. Instead, it would seem more parsimonious to suggest that both visual attributes to be bound are capable of gating the activity of their V1 afferents through feedback, but that the dynamics of this feedback process vary between attributes. The previous sections have ignored the dynamics of chromatic processing, implicitly assuming its effect on perceptual asynchrony is negligible compared to the effects of motion and orientation. Whether this assumption is ultimately valid is a question that should be amenable to physiological investigation.

13.11 Conclusion

In this chapter I have discussed the relationship between the dynamics of visual feature binding and the time course of the corresponding neural processes. Psychophysical studies of perceptual binding between pairs of colors and directions of motion reveal effects of angle of direction change and motion transparency that can be related to opponent motion inhibition. Further evidence of a strong link between the dynamics of visual processing and the time course of visual perception comes from work using contingent adaptation as a metric of interattribute binding, complementing studies relying exclusively on subjective perceptual report. Diagnostic to the mechanisms of perceptual binding is the observation that color and orientation can be correctly bound at high rates of temporal alternation even though a systematic temporal bias in their binding is evident at longer periods. A mechanistic account of the dynamics of binding color with direction of motion and orientation is proposed that ascribes a central role to feedback from extrastriate areas to early visual cortex. This account relies upon the high temporal fidelity of responses in early visual cortical areas to enable high temporal resolution binding. The systematic temporal biases observed at longer alternation periods are attributed to temporal aspects of processing in extrastriate areas. The interaction of these areas in dynamically binding visual features is proposed to be mediated by feedback from the relatively modular extrastriate areas to neurons in early visual cortex that encode multiple attributes in combination. Within this framework, the experienced sequence of events is not simply determined by the latencies of bottom-up perceptual processing but is also influenced by the dynamics of feedback from higher-level interpretations of the stimulus in terms of surface representations.

Acknowledgments

The author is supported by an Australian Research Fellowship and Discovery Project from the Australian Research Council. Thanks to Alex Holcombe, Erin Weston, Derek Arnold, Szonya Durant, and Joel Pearson for helpful discussions and to Beena Khurana and two anonymous reviewers for constructive criticism.

References

Albright, T. D., & Desimone, R. (1987). Local precision of visuotopic organization in the middle temporal area (MT) of the macaque. *Exp Brain Res* **65**: 582–592.

Arnold, D. H. (2005). Perceptual pairing of colour and motion. *Vision Res* **45**: 3015–3026.

Arnold, D. H., & Clifford, C. W. G. (2002). Determinants of asynchronous processing in vision. *Proc R Soc Lond B* **269**: 579–583.

Arnold, D. H., Clifford, C. W. G., & Wenderoth, P. (2001). Asynchronous processing in vision: colour leads motion. *Curr Biol* **11**: 596–600.

Barbur, J. L., Wolf, J., & Lennie, P. (1998). Visual processing levels revealed by response latencies to changes in different visual attributes. *Proc R Soc Lond B* **265**: 2321–2325.

Bartels, A., & Zeki, S. (1998). The theory of multistage integration in the visual brain. *Proc R Soc Lond* B **265**: 2332–2337.

Bedell, H. E., Chang, S. T. L., Ogmen, H., & Patel, S. S. (2003). Color and motion: which is the tortoise and which is the hare? *Vision Res* **43**: 2403–2412.

Blake, R., & He, S. (2005). Adaptation as a tool for probing the neural correlates of visual awareness: progress and precautions. In: C. W. G. Clifford & G. Rhodes (eds.), *Fitting the Mind to the World: Aftereffects in High-Level Vision* (281–307). Oxford: Oxford University Press.

Blaser, E., Papathomas, T., & Vidnyanszky, Z. (2005). Binding of motion and colour is early and automatic. *Eur J Neurosci* **21**: 2040–2044.

Clifford, C. W. G. (2005). Functional ideas about adaptation applied to spatial and motion vision. In: C. W. G. Clifford & G. Rhodes (eds.), *Fitting the Mind to the World: Aftereffects in High-Level Vision* (47–82). Oxford: Oxford University Press.

Clifford, C. W. G., Arnold, D. H., & Pearson, J. (2003). A paradox of temporal perception revealed by a stimulus oscillating in colour and orientation. *Vision Res* **43**: 2245–2253.

Clifford, C. W. G., Holcombe, A. O., & Pearson, J. (2004). Rapid global form binding with loss of associated colors. *J Vis* **4**: 1090–1101.

Clifford, C. W. G., Spehar, B., & Pearson, J. (2004). Motion transparency promotes synchronous perceptual binding. *Vision Res* **44**: 3073–3080.

Dennett, D. C., & Kinsbourne, M. (1992). Time and the observer: the where and when of consciousness in the brain. *Behav Brain Sci* **15**: 183–247.

Eagleman, D., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science* **287**: 2036–2038.

Enns, J. T., & Oriet, C. (2004). Perceptual modularity: modularity of consciousness or object updating? *J Vis* **4**: 27a.

Favreau, O. E., Emerson, V. F., & Corballis, M. C. (1972). Motion perception: a color-contingent aftereffect. *Science* **176**: 78–79.

Forte, J., & Clifford, C. W. G. (2005). Interocular transfer of the tilt illusion shows that monocular orientation mechanisms are colour selective. *Vision Res* **45**: 2715–2721.

Frisby, J. P. (1980). *Seeing: Illusion, Brain and Mind*. Oxford: Oxford University Press.

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature* **383**: 334–337.

He, S., & MacLeod, D. I. (2001). Orientation-selective adaptation and tilt after-effect from invisible patterns. *Nature* **411**: 473–476.

Hochstein, S., & Ahissar, M. (2002). View from the top: hierarchies and reverse hierarchies in the visual system. *Neuron* **36**: 791–804.

Holcombe, A. O. (2001). A purely temporal transparency mechanism in the visual system. *Perception* **30**: 1311–1320.

Holcombe, A. O., & Cavanagh, P. (2001). Early binding of feature pairs for visual perception. *Nat Neurosci* **4**: 127–128.

Horwitz, G. D., & Albright, T. D. (2005). Paucity of chromatic linear motion detectors in macaque V1. *J Vis* **5**: 525–533.

Humphrey, G. K., & Goodale, M. A. (1998). Probing unconscious visual processing with the McCollough effect. *Conscious Cogn* **7**: 494–519.

James, W. (1890). *The Principles of Psychology, Vol. 1*. New York: Henry Holt.

Jeannerod, M. (1992). The where in the brain determines the when in the mind. *Behav Brain Sci* **15**: 212–213.

Johnston, A., & Nishida, S. (2001). Time perception: brain time or event time? *Curr Biol* **11**: R427–R430.

Johnson, E. N., Hawken, M. J., & Shapley, R. (2001). The spatial transformation of color in the primary visual cortex of the macaque monkey. *Nat Neurosci* **4**: 409–416.

Kanai, R., Paffen, C. L. E., Gerbino, W., & Verstraten, F. A. J. (2004). Blindness to inconsistent local signals in motion transparency from oscillating dots. *Vision Res* **44**: 2207–2212.

Mather, G., Verstraten, F., & Anstis, S. (eds.). (1998). *The Motion Aftereffect: A Modern Perspective*. Cambridge, MA: MIT Press.

Mayhew, J. E. W., & Anstis, S. M. (1972). Movement aftereffects contingent on colour, intensity and pattern. *Perception & Psychophysics* **12**: 77–85.

Moradi, F., & Shimojo, S. (2004). Perceptual-binding and persistent surface segregation. *Vision Res* **44**: 2885–2899.

Moutoussis, K., & Zeki, S. (1997a). A direct demonstration of perceptual asynchrony in vision. *Proc R Soc Lond* B **264**: 393–399.

Moutoussis, K., & Zeki, S. (1997b). Functional segregation and temporal hierarchy of the visual perceptive systems. *Proc R Soc Lond B* **264**: 1407–1414.

Munk, M. H. J., Nowak, L. G., Girard, P., Chounlamountri, N., & Bullier, J. (1995). Visual latencies in cytochrome oxidase bands of macaque area V2. *Proc Natl Acad Sci USA* **92**: 988–992.

Nishida, S., & Johnston, A. (2002). Marker correspondence, not processing latency, determines temporal binding of visual attributes. *Curr Biol* **12**: 359–368.

Pascual-Leone, A., & Walsh, V. (2001). Fast back projections from the motion to the primary visual area necessary for visual awareness. *Science* **292**: 510–512.

Qian, N., & Andersen, R. A. (1994). Transparent motion perception as detection of unbalanced motion signals. II. Physiology. *J Neurosci* **14**: 7367–7380.

Qian, N., Andersen, R. A., & Adelson, E. H. (1994a). Transparent motion perception as detection of unbalanced motion signals. I. Psychophysics. *J Neurosci* **14**: 7357–7366.

Qian, N., Andersen, R. A., & Adelson, E. H. (1994b). Transparent motion perception as detection of unbalanced motion signals. III. Modeling. *J Neurosci* **14**: 7381–7392.

Rajimehr, R. (2004). Unconscious orientation processing. *Neuron* **41**: 663–673.

Schiller, P. H., & Malpeli, J. G. (1978). Composition of geniculostriate input to superior colliculus of the rhesus monkey. *J Neurophysiol* **41**: 788–797.

Seidemann, E., Poirson, A. B., Wandell, B. A., & Newsome, W. T. (1999). Color signals in area MT of the macaque monkey. *Neuron* **24**: 911–917.

Shadlen, M. N., & Newsome, W. T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J Neurosci* **18**: 3870–3896.

Shady, S., MacLeod, D. I., & Fisher, H. S. (2004). Adaptation from invisible flicker. *Proc Natl Acad Sci USA* **101**: 5170–5173.

Shipp, S., & Zeki, S. (1989). The organization of connections between areas V5 and V1 in macaque monkey visual cortex. *Eur J Neurosci* **1**: 309–332.

Snowden, R. J., Treue, S., Erickson, R. G., & Andersen, R. A. (1991). The response of area MT and V1 neurons to transparent motion. *J Neurosci* **11**: 2768–2785.

van Doorn, A. J., & Koenderink, J. J. (1982). Temporal properties of the visual detectability of moving spatial white noise. *Exp Brain Res* **45**: 179–188.

Viviani, P., & Aymoz, C. (2001). Colour, form, and movement are not perceived simultaneously. *Vision Res* **41**: 2909–2918.

Wu, D. A., Kanai, R., & Shimojo, S. (2004). Vision: steady-state misbinding of colour and motion. *Nature* **429**: 262.

Zeki, S. (1993). *A Vision of the Brain.* Oxford: Blackwell.

Zeki, S. (2003). The disunity of consciousness. *Trends Cogn Sci* **7**: 214–218.

Zeki, S., & Bartels, A. (1998). The asynchrony of consciousness. *Proc R Soc Lond B* **265**: 1583–1585.

Zeki, S., & Bartels, A. (1999). Toward a theory of visual consciousness. *Conscious Cogn* **8**: 225–229.

14

How does the timing of neural signals map onto the timing of perception?

DAVID M. EAGLEMAN

Summary

Different features of stimuli are processed in the central nervous system at different speeds. However, such neural time differences do not map directly onto perceptual time differences. How the brain accounts for timing disparities to correctly judge the temporal order of events in the world is the temporal binding problem. I weigh physiological data against new psychophysical findings both within and between modalities. The essence of the paradox is that the timing of neural signals appears, at first blush, too variable for the high accuracy of the psychophysical judgments. I marshal data indicating that ~80 msec is an important duration in perception and make the novel suggestion that this number is directly mirrored in the physiology. In recordings from several areas of the primate visual system, the difference between the slowest and fastest latencies based on luminance contrast is 80 msec. If the rest of the brain wants to time outside stimuli correctly, it must account for the fact that the earliest stages of the visual system spread signals out in time. I suggest that the brain waits for the slowest information to arrive before committing to a percept. This strategy only applies to visual awareness; in contrast, the motor system may form its reactions based on the first incoming spikes.

14.1 Introduction

One goal of modern neuroscience is to relate physiological data to perception (Eagleman 2001). How do spikes recorded from single neurons map onto object recognition, brightness perception, or timing judgments? Despite decades of work, there are few good theories uniting the wetware to the perception, and we currently cannot build machines that are visually aware. This chapter attempts to ferret out some relationships between the two domains that can guide our search. My strategy is to come at the problem from the point of view of time. If we can find parallels in the temporal relationships between the physiology and the psychophysics, this may open new inroads into their interaction.

Nervous systems face the problem of feature binding – that is, keeping features of an object perceptually united such that, for example, the redness and squareness do not bleed off a moving red square (Crick & Koch 1990; Engel et al. 1992). The fact that feature binding is usually performed correctly would not come as such a surprise were it not for our modern picture of the mammalian brain, in which different kinds of information are processed in

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

different neural streams. Binding requires coordination not only among different modalities (vision, audition, olfaction, etc.) but also among different features within a modality (in vision, e.g., color, motion, form identification).

But there is a deeper challenge the brain must tackle, without which feature binding would rarely be possible. This is the problem of *temporal* binding: the assignment of the correct timing of events in the world. The challenge here is that different stimulus features move through different processing streams – and *are processed at different speeds*. I will draw on electrophysiological data that follow to demonstrate this point. The brain must account for speed disparities between and within its different sensory channels if it is to accurately determine the timing relationships of features in the world.

Many discussions of neural function tacitly rest on the assumption that awareness (what the subject reports) is an online phenomenon, coming about as soon as the leading edge of the represented stimulus reaches a "perceptual end-point" (Zeki & Bartels 1998a). However, many demonstrations in the literature suggest that awareness is not an online phenomenon but is instead delayed (Kolers & von Grunau 1976; Dennett 1991; Kinsbourne 1993; Bachmann 1994; Pessoa et al. 1998; Eagleman & Sejnowski 2000a,b,c, 2003, 2007; Eagleman 2008). That is, perceptions are retrospectively attributed after the brain has gathered information from a window of time around an event. I will argue here that this postdictive framework is a necessary component to any theory of awareness and of its necessity to solve the temporal binding problem. I will then discuss what this illuminates about the physiology.

In this chapter I draw mainly on data from the visual system, but the framework proposed here applies to all modalities, and I will give examples where available. I will sketch the relevant neurophysiology of the visual system, illustrating that signals reflecting different stimulus features move through the system at different times. This leads us to ask how visual perception ever gets the timing correct. I will ask why some optical illusions exist and why others do not. Finally, I will suggest that temporal binding can be accomplished via a window of delay that allows for more slowly processed information to participate in the interpretation. In other words, the brain waits to collect the slowest signals. I draw on neurophysiologic measures to suggest that this window should be about 80 msec. This window of delay means that awareness is postdictive, incorporating data from a window of time both before and after the event and delivering a retrospective *interpretation* of what happened (Eagleman & Sejnowski 2000a,b,c; Eagleman 2008). Postdiction is the opposite of prediction, which is the act of extrapolating into the future to guess about events that have not yet happened. I conclude with a discussion of how signals can be delayed and aligned in neural tissue.

14.2 Physiology of the visual cortex

A measure of increasing importance to physiologists is the *latency* of a neuron's response: how much time passes between the onset of a stimulus and a cell's first measurable response to it. Figure 14.1 shows that latencies are surprisingly variable across different parts of the

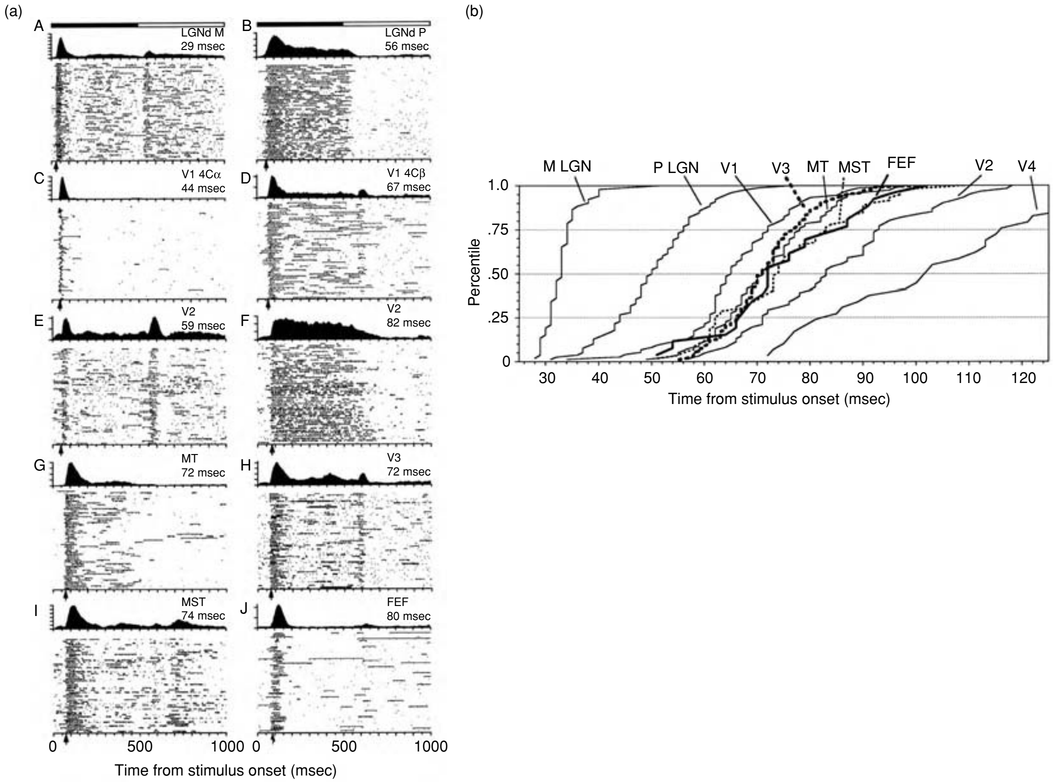


Fig. 14.1 Signal timing across the macaque visual system; data from Schmolesky et al. (1998). There is a wide spread of latencies both between and within areas. How does the visual system get the timing of events correct in the face of this signal variability? (a) Spikes are displayed as ticks in the raster panels, with fitted spike density histograms above. Black bar indicates the time the stimulus was on (0–500 msec). Arrow beneath each raster plot indicates the visual response latency of the cell; the value is also displayed in the top right corner. (b) Cumulative distributions of latencies. Percentage of cells that have begun to respond is plotted as a function of time from stimulus presentation.

brain – in this case, across different areas of the visual system (Schmolesky et al. 1998). When a stimulus appears in the outside world, the response in the brain is smeared out over a large window. Think of Paul Revere and his colleagues spreading out in different directions over the New England landscape to deliver a message. Some riders are fast, others slow. Almost all of them inspire other riders to saddle up and ride off in different directions. As a result, the colonialists do not all get the message at once; it percolates at different rates to different streets in different townships. As can be seen in Fig. 14.1, there is also a good deal of trial-to-trial variability within a single neuron: in other words, each time you rerun history, each horseman may ride faster or slower than the last time.

What is mysterious about this physiology is the fact that humans have quite good resolution when making temporal judgments. For example, two visual stimuli can be accurately deemed simultaneous to 5-msec resolution, and their order can be assessed with 20-msec resolution (Hirsh & Sherrick 1961). How do the colonies conclude exactly when the British arrived, given the spread of signals in different locations? I will address this paradox by focusing on a well-studied example – the effect of intensity on latency – to drill down to the bottom of the mystery.

14.3 Contrast differences lead to latency differences

Latencies can be modulated in a specific reliable manner by changing the intensity of a stimulus. For example, in a study by Maunsell and colleagues, a monkey sat in a dark room and was sporadically presented with a flash (Maunsell et al. 1999). The flash ranged in luminance from low (0.43 cd/m^2) to high (28 cd/m^2). Figure 14.2(a) shows the response of cells in dorsal lateral geniculate nucleus (dLGN), a midpoint between the retina and visual cortex. As can be seen from the figure, low-luminance stimuli cause smaller, and *later*, responses than higher luminance stimuli. This conclusion holds irrespective of the measure of latency (time to onset, half-peak, or peak; Fig. 14.2(b)). In other words, at the early stages of the visual system, even before reaching visual cortex, signals are already becoming spread out through time based on stimulus properties.

This fact is true in primary visual cortex as well. The contrast of an oriented bar changes the response latency of neurons in V1 (Fig. 14.2(c)) (Gawne et al. 1996). In higher visual areas, such as STS, stimulus contrast is the major factor affecting response latency, independent of the response magnitude (Oram et al. 2002).

These data raise a critical question: What does this temporal spread based on intensity mean for perception? Let’s look at a few possibilities.

14.4 The online hypothesis

The first possibility we will consider is a popular, enduring, and likely incorrect view. The view is that neural latency differences between two stimuli will translate into perceptual time differences (Purushothaman et al. 1998; Whitney & Murakami 1998; Zeki & Bartels 1998b; Patel et al. 2000; Whitney & Cavanagh 2000). I will refer to this as the *latency*

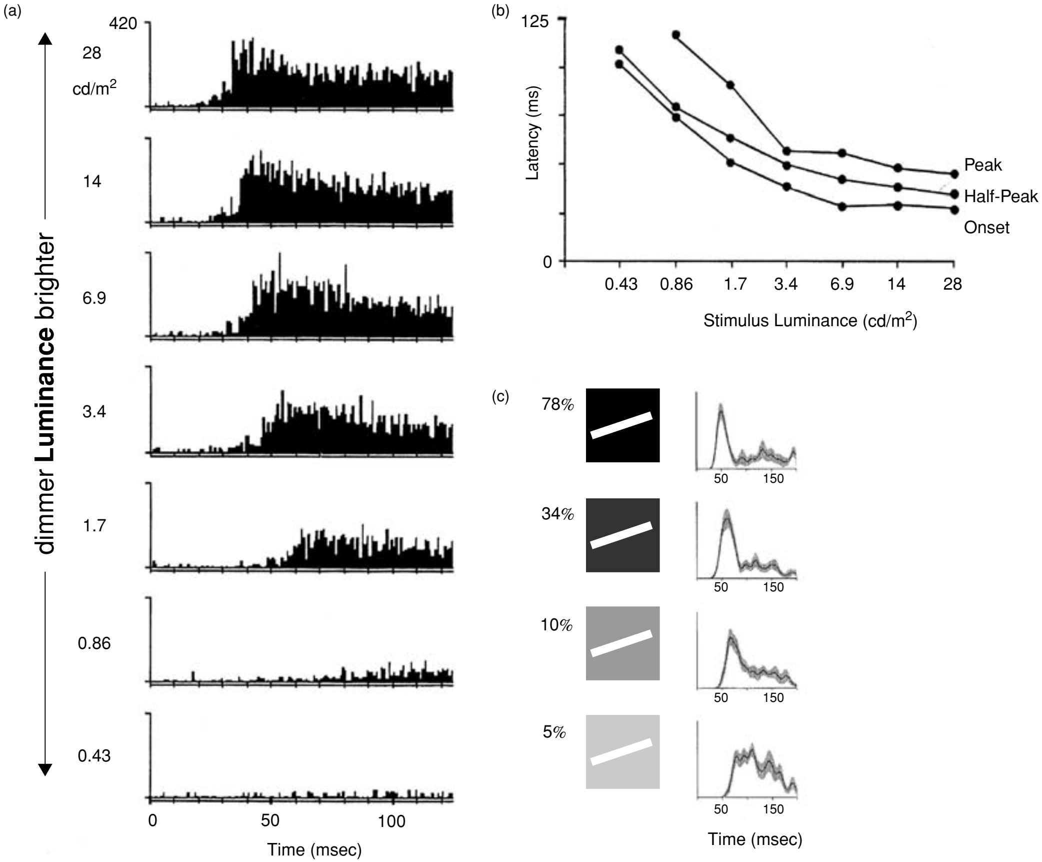


Fig. 14.2 Latency as a function of stimulus intensity. (a) Peristimulus time histograms of single neurons in dLGN in response to flashes of different luminance. Data from Maunsell et al. (1999). (b) The same data plotted differently to emphasize that latencies become shorter as the stimulus gets more intense, irrespective of the method for measuring those latencies. (c) Data from Gawne et al. (1996). Cells in V1 respond with later latency as the contrast of the stimulus is reduced.

difference or *online* hypothesis. Essentially, this position holds that differences in neural latencies (as measured, say, in LGN or cortex, Fig. 14.2) map directly onto temporal differences in perception. Take as an example two simultaneously appearing stimuli, X and Y, each with different features. If stimulus X causes a cortical response before Y, the online hypothesis states that it is perceived first. The online hypothesis immediately raises some tricky questions (such as where, exactly, is the endpoint at which the rabbit and turtle are compared?). But we'll overlook those for the moment.

Because the online hypothesis is a seductive and pervasive habit of thinking, let's take it seriously for a moment. If it were true that stimuli were perceived online, that is, as soon as the leading edge of information reached some finish line in visual cortex, then neural correlates of perception would be easy to measure: whatever arrives in visual cortex first is perceived first. But some simple thought experiments call the theory into question. We will entertain the online hypothesis, show where it breaks down, and then take a look at its alternatives.

First, if there were illusions from differential latencies from onset, we should expect to suffer motion illusions each time we blink our eyes or turn on the lights. Following either of these events, everything in the visual field has a simultaneous onset. Given that there are bright and dim parts in the visual scene, wouldn't we see illusory movement – the bright areas of the room apparently moving toward the dim bits? It can be easily demonstrated for oneself that this does not occur.

Next, present to yourself a series of photographs flashed rapidly in sequence: a house, a tiger, a car, and so on. Even though each picture contains regions of low and high contrast, it is rarely temporally confused – that is, one does not generally perceive the stripes of the tiger on the house, the headlights on the tiger, and so on. Note that Intraub (1985) reported a temporal dissociation during rapid serial visual presentation, but only for a frame surrounding an outline drawing, and never for the contents of the drawing itself.

To give a more traditional psychophysical example, we turn to a series of optical illusions that should exist but do not. To directly address the online hypothesis, I have engineered a series of simple experiments, shown in Fig. 14.3. First, consider a vertical stack of squares with differing contrasts that move (as a unit) horizontally back and forth across the field of view (Fig. 14.3, top left). The online hypothesis predicts that, because the brightest square enjoys the lowest latency, it may be perceived faster, whereas the dimmest square is getting processed most slowly (and so on for the squares in between). Even though the delay is constant, one might expect that the brightest square is always perceived at a position well ahead of the dimmest, and the colinear squares may appear to become noncolinear. However, all observers tested report that no such illusion occurs ($n = 8$). A similarly negative result is obtained with spinning bars with gradient textures (Fig. 14.3, bottom left). At all speeds tested, the bars appear solid, not curved or rubbery.

The theme is repeated with horizontally moving gradient squares (Fig. 14.3, top right). An online model predicts the contraction or dilation of the squares, depending on their direction of movement. That is, if the higher-contrast parts of the square have shorter latencies (and thus faster perception, in the online view), then a bright leading edge and a dim lagging edge

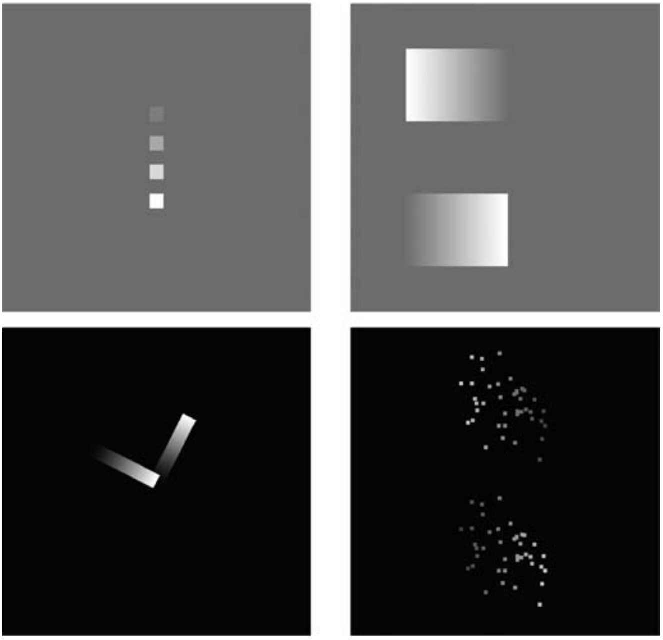


Fig. 14.3 Optical illusions that should exist but do not. Top left: A tower of four squares of differing contrast moves horizontally back and forth. An "online" interpretation of the physiology in Fig. 14.2 would predict that higher contrast squares should appear to be ahead of lower contrast squares. Instead, all observers report no illusion. The squares look vertically aligned, as they should be. Bottom left: Bars with gradients spin clockwise. An online hypothesis would predict that the bar does not look rigid. Again, the expected illusion does not occur. Top right: Two squares with contrast gradients move back and forth horizontally. An online hypothesis would predict that the square moving in the direction of its high luminance should look wider than the other square that moves in the direction of low luminance, which should look narrower. The illusion does not obtain. Bottom right: Same as above, but this time with a moving swarm of flashes. The flashes appear with a gradient of luminances. As above, the online hypothesis predicts a shrinking or stretching of the distribution in the direction of its motion, yet no such illusion occurs.

may appear to stretch out the distribution. Moving in the opposite direction, the square may appear to contract. I constructed a demonstration using two squares with opposite gradients for direct comparison to each other. As the squares moved back and forth, all observers reported there is no perceptual illusion of stretching or shrinking. To rule out the possibility that the deformation of the square was too small, I artificially stretched or constricted the squares by the amount predicted by the physiology in Fig. 14.2(a), using the difference between bright and dim signals (the details of this choice will be justified later). Subjects were 100% accurate at detecting the artificially stretched or shrunken squares.

Perhaps one will argue that the edges of the squares give extra information. So in the final demonstration, the moving squares were replaced with random dot distributions (Fig. 14.3, bottom right). The dots were assigned luminances based on a horizontal gradient from low to high. The dots flashed on and off, never appearing in the same relationship to one another, and the distribution as a whole moved horizontally back and forth. Again, the distributions did not appear to stretch and shrink in relation to one another, as the online hypothesis would have erroneously predicted. An artificially shrinking or stretching distribution was easily detectable (see online demonstration).

Whatever other effects one might argue are occurring here (e.g., motion capture), the conclusion is the same: the visual system gets the frame-by-frame timing right (see also Kopinska et al. 2003, for similar results). Below I will suggest how it does so.

14.5 Reconciling latency differences with perceptual simultaneity, or, the magic 80 msec: waiting for the slowest signals

The experiments illustrated earlier that even though stimuli of different luminances give rise to widely different response latencies at the first stages of the visual system, these do not have a direct mapping onto differences in perceptual time. So how are these latencies reconciled?

I suggest that the answer comes from looking at the tasks and resources of the visual system. As one of its tasks, the visual system tries to get the timing of outside events correct. But for its resources, it has to deal with the peculiarities of the equipment that feeds it: the eyes and the thalami. These have their own evolutionary histories and idiosyncratic circuitries – and because of the details of their wiring, signals become spread out in time from the first stages of the visual system (e.g., Maunsell et al. 1999).

So if the visual brain, the recipient of smeared temporal information, wants to get events correct, time-wise, it may have only one choice: *wait for the slowest information to arrive*.

How long would the system have to wait? According to the physiology, I suggest it would have to wait about 80 msec to collect all the information, from the dimmest to the brightest. This number can be read directly from Fig. 14.2(a) and (b): the latency difference between the dimmest and brightest stimuli is 80 msec on the y-axis.

Note that the ~80-msec time window (plus or minus 20 msec) crops up commonly in psychophysics. For example, this window is found in motion integration, motion deblurring (Burr & Morgan 1997), successive pattern integration (Di Lollo 1980), binocular pattern integration (Julesz & White 1969; Ross & Hogben 1974), backward masking (Bachmann 1994; Macknik & Livingstone 1998), and audio–video synchronization (Steinmetz & Engler 1993). In our previous work on the flash-lag effect, we showed an approximately 80-msec window over which a moving object could be manipulated *after* a flash and still achieve an effect on what the viewer reports having seen at the time of the flash (Eagleman & Sejnowski 2000a,b,c).

To my knowledge, this is the first suggestion that the psychophysical 80-msec window is directly mirrored in the physiology. That is, if the brain wants to wait for the slowest information, it must wait about 80 msec. This would allow the visual system to discount latency differences imposed by the early stages, but it has the disadvantage of pushing perception into the past. Counterbalancing the need to collect slow information is the survival advantage to operating as close to the present as possible. In other words, a system would not want to live too far in the past. I suggest that 80 msec is the smallest delay that allows higher areas of the brain to account for the latencies engendered by the first stages of the system while still operating close to the present. Note that by studying the psychophysics of simultaneity across modalities, Kopinska and Harris (2004) also concluded that the brain waits for the slowest information to arrive – in their case, they suggested a delay of 94 msec.

Among other things, this strategy of waiting for the slowest information has the great advantage of allowing object recognition to be *independent of lighting conditions*. Imagine a striped tiger coming toward you under the canopy of a forest, passing through different patches of sunlight. Neuroscience currently appreciates how difficult the task of object recognition is for the visual system. Now imagine how much harder the task would be if the shadow-play across the tiger caused incoming signals to be processed at different speeds. Different fragments of the tiger would sunder into different locations. Somehow the visual system has evolved to reconcile different information latencies; after all, it is advantageous to recognize tigers regardless of the lighting.

We should note that it has been suggested that latencies could be used by the system as a code (Oram et al. 2002), perhaps for object recognition (Thorpe et al. 2001) or object feature binding (Gawne et al. 1996). For example, if signals arriving along transmission lines arrive in the order 3–6–2–4, that could code for ostrich, whereas 2–7–4–5 codes for carrot. This idea is known as rank order coding. Despite the appeal of its simplicity, I suggest that using latencies for object recognition is untenable in biology – and this is because the luminances from an object are dependent on the lighting conditions. If you learn to recognize a statue of George Washington in the morning, the lighting angle in the afternoon will make the contrast-based latencies entirely different – and a rank order coding network will utterly fail to recognize it. Marrying one's visual recognition capacities to the particulars of the lighting conditions is not a move Mother Nature seems to have taken.

Finally, an important point needs to be clarified. The 80-msec window under discussion does not imply that the visual system sees in "chunks." The idea that the visual system takes discrete "snapshots" has been entertained (Varela et al. 1981; VanRullen & Koch 2003), but the available evidence speaks against it (Kline et al. 2004; Kline & Eagleman 2008). Instead, the 80-msec window we are discussing appears to be a duration over which the visual system waits to collect information, but it can still retain the ability to differentiate events within that window (Blake & Lee 2005). For example, if I were to collect the dot-dash-dot-dot that designates the letter "L" in Morse code, I can know that the temporally smeared information applies to one object and yet still report on the order of dots and dashes. Further, I speculate that an 80-msec window may only be triggered when a perceptual question is asked; this is the topic of future exploration (Eagleman & Dennett, in preparation).

14.6 Latency illusions that do exist, and why

Getting back to the examples in Fig. 14.3, these may strike the reader as a contradiction of other reports in the literature. For instance, in the Pulfrich effect, a pendulum appears to rotate in depth when a neutral density filter is placed in front of one eye. Following a

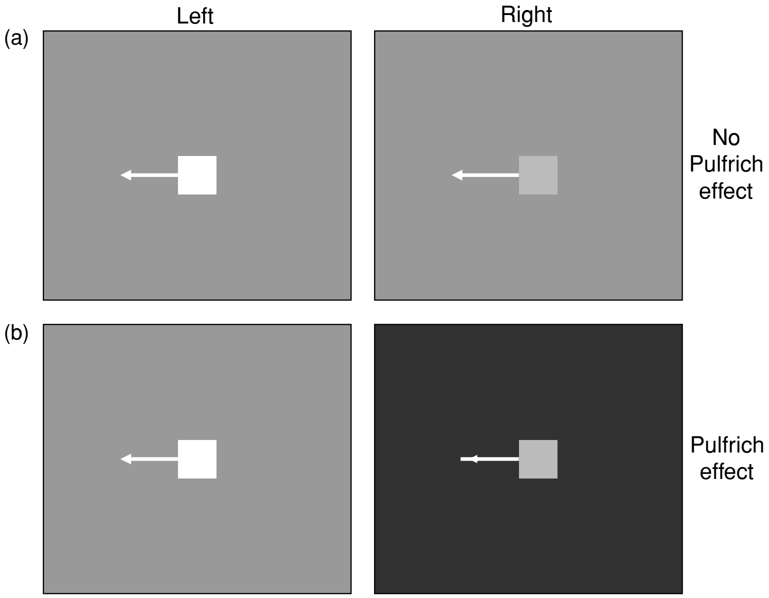


Fig. 14.4 Conditions for the Pulfrich effect. (a) The Pulfrich effect is not obtained simply by reducing the contrast of the moving stimulus to one eye (in this case, the square presented to the right eye). (b) Instead, the effect is only obtained when the presentation to one eye is globally darkened, as with a neutral density filter. We suggest that this is because the visual system has evolved to deal with a range of luminance differences in normal vision but has not evolved to deal with the pathologic case of different global levels of illumination between the two eyes.

suggestion by Fertsch, Pulfrich hypothesized a timing difference between signals from the two eyes (Pulfrich 1922; Eagleman 2001). We have every reason to believe that latency differences are the best explanation for the Pulfrich effect, so this is often cited in support of an online model. However, it is critical to note that the Pulfrich effect is a special case. Specifically, our argument is that the visual system has evolved to account for latency differences such that under *normal circumstances* it will not be fooled by variations in luminance. The Pulfrich effect is a totally unnatural stimulus because one retina is reading information at an entirely different mean-luminance level than the other retina. The impulse response function on the dimmer retina is greatly slowed (Purpura et al. 1990), leading to the illusion. Of the many assumptions that the brain makes, an important one is that both eyes are viewing the same luminance in the scene. The visual system has not evolved to deal with deviations from this condition, and it is not surprising that an illusion can be induced.

In support of this argument, I have discovered a fact that appears not to be widely known. When observers are shown a binocular presentation of a square that moves back and forth horizontally, the Pulfrich illusion is *not* produced when one merely lowers the luminance of one of the moving squares (Fig. 14.4(a)). This negative result tends to come as a surprise to psychophysicists familiar with the illusion. Instead, the effect is produced only when the entire display to one eye is darkened, as with a neutral density filter (Fig. 14.4(b)). In other words, the effect of a neutral density filter over one eye is not simply to reduce the luminance of the moving object but more broadly to expose the entire retina to a different mean luminance.

I have found the same to be true of the Hess effect, an illusion in which an offset can be perceived between two horizontally moving dots when one is of high contrast and the other low contrast (Wilson & Anstis 1969). Specifically, I find that this effect is only seen when one uses a neutral density filter over half the screen – simply reducing the contrast of a single dot is insufficient (at least under phototopic viewing conditions). Fooling the visual system with a latency difference requires slowing the signals through all or part of the retina – simply changing the luminance of the moving stimulus is insufficient because, as I have argued earlier, the visual system is equipped to deal with and account for different object luminances in normal vision.

14.7 Temporally spread signals in the neural tissue

So we have asserted that the brain can keep account of latencies. But how exactly could it know what happened when? To highlight this problem, we can phrase it as a question: If you were a V1 neuron and received a burst of spikes, how would you know if that meant a dim flash occurred 150 msec ago or a bright flash occurred 70 msec ago? There may be at least two strategies the system can employ to take care of this problem with temporally spread, delayed signals: "timestamp" them or physically correct them on the fly (i.e., temporally align the signals). Because we do not know which the system does, we will briefly sketch out possible methods for both.

First, I'll address the timestamp model. Even at the single-neuron level, there is more information available than simply the latency. Notice in Fig. 14.2(a) that the different spike trains have different temporal signatures. For instance, the spike rate at 28 cd/m^2 has a sharper onset than the spike train at 1.7 cd/m^2. These structures in the spike timing (or the interspike interval) could in theory allow downstream neurons to distinguish an old dim flash from a recent bright flash merely by the structure through time. This would be one way to reconstruct the actual order of events.

As an alternative, neural circuitry could reconcile differential latencies by physically aligning the timing of signals. Although most of the available data in the field comes from single electrode recordings, it is critical to keep in mind that *populations* of neurons could manipulate both latencies and the variability of those latencies. For example, one way the brain could align signals is by dynamically recruiting more or fewer neurons to speed and slow the passage of signals. For example, a weaker signal (lower spike rate) could recruit more neurons downstream, which could bring further neurons to threshold more quickly. Conversely, faster spike rates activate fewer downstream neurons, such that the timing is slowed. The mechanism for accomplishing this could be fast synaptic depression, with the result that the timing of simultaneous events in the world would be temporally "lined up" at higher stages of the nervous system. In fact, as pointed out by Maunsell et al. (1999), the faster speed of the magnocellular pathways cannot be predicted solely from differences in axon conduction speeds; instead the degree of convergence may be highly regulatory in timing issues. For instance, different degrees of convergence (beginning with retinal circuits and continuing through the brain), can speed or retard timing – such that a weaker signal in a highly convergent pathway might lead to faster downstream responses than a faster signal in a pathway with less convergence (Maunsell et al. 1999). So although the details of circuitry are most commonly thought of in terms of connectionist functions (Callaway 1998), the circuitry could also, in theory, lead to changes in timing. If true, this could in the future inspire us to rethink circuitry not only in terms of spatial connections but also what they are implementing temporally.

Although some possibilities for speeding and slowing signals have been sketched, it is critical to remember that there is no theoretical necessity for neural signals to line up temporally for the perception of simultaneity. As in an earlier example, the letter "L" is represented in Morse code by a dot, a dash, and two more dots. When someone receives this temporal signal, it is interpreted as an "L." The same could hold for perception: signals arriving at different times could be interpreted as simultaneous events in the outside world. In other cases, simultaneously arriving signals might be best interpreted by perception as asynchronous events in the real world. As Uttal stated this point: "The essence of much of the research that has been carried out in the field of sensory coding can be distilled into a single, especially important idea – any candidate code can represent any perceptual dimension; there is no need for an isomorphic relation between the neural and psychophysical data. Space can represent time, time can represent space, place can represent quality, and certainly, nonlinear neural functions can represent linear or nonlinear psychophysical functions equally well" (Uttal 1979).

Finally, we note the importance of massive feedback connectivity in visual awareness (Mumford 1994; Nowak & Bullier 1997; Lamme & Roelfsema 2000). It is in this light that "waiting for the slowest signal" must be finally understood physiologically. Our intuition is that the 80-msec window will not be found in terms of a neural information buffer in a feedforward buffer, but rather in the settling of recurrent networks into a larger pattern. We do not at present know the details of how this works; this is exciting open ground for the future.

14.8 Conclusions

The problem faced by the visual system is a common problem for biological creatures in a temporal world. Understanding the timing of events is critically important, but the signals representing that timing may be spread out in time. This requires a reconstruction of event timing, whether implicitly (symbolic coding) or explicitly (aligning signals in time). Current data are too sparse to arbitrate between these two methods; the nervous system may use either or both.

I have argued that the back of the brain seeks to judge timing of events accurately, but it has to contend with the temporally smeared information sent to it by the eyes and thalamus. The best solution to this problem may be to wait for the slowest information to arrive. Electrophysiology from the primate visual system shows that the window over which information should be collected is about 80 msec, the latency difference between the fastest and slowest signals. I suggest this is the physiological basis of the 80-msec window observed in a large variety of psychophysics experiments, as detailed earlier.

Among other things, collecting signals over a window of time allows object recognition independent from lighting conditions, because latency differences based on different luminances can be discounted. In contrast, models that tie themselves to latency differences for object recognition (Thorpe et al. 2001) are unfortunately tied to the details of the lighting, a disadvantageous move for any visual organism.

The argument I have made – that the system waits to collect information over the expected window of time over which it will come streaming in – applies not only in vision, but more generally to all modalities. Therefore, although it is possible to measure an 80-msec window of postdiction in vision (Eagleman & Sejnowski 2000), the breadth of this window may be different in audition and somatosensation. It may therefore be that a unified polymodal perception of the world has to wait for the slowest overall modality. Given conduction times along limbs, this leads to the bizarre but testable suggestion that tall people may live further in the past than short people. The consequence of waiting for temporally spread signals is that perception becomes something like the airing of a "live" television show. Such shows are not truly aired live but are instead delayed by a small window of time in case editing becomes necessary.

Note that the 80-msec window proposed here is the minimum duration that the visual system would have to wait to collect all the information from a visual event. However, it is possible that when including signal travel time and multimodal unification, the total lag between a physical moment and its conscious perception is much longer. Although the total delay between a stimulus and conscious awareness has been proposed to be impossible to know for certain (Dennett 1991), some investigators have estimated the total time-to-awareness in the range of 100–150 msec (Lamme 2003) to 500 msec (Libet et al. 1967).

It must be emphasized that everything I have discussed in this chapter is about visual awareness. It seems clear from preconscious reactions that the motor system does not wait for all the information to arrive before making its decisions. In general, the motor systems can act appropriately with partial or no participation of awareness (Goodale & Milner 1992, 2004) – for example, visual information streams directly to the amygdala, which may have a direct and rapid connection to the motor systems.

This raises a question: What is the *use* of perception, especially given our argument that perception lags reality, is only retrospectively attributed, and is generally outstripped by automatic (unconscious) systems? The most likely answer is that perceptions are representations of information that the brain can manipulate at a later date – these representations can be worked with by cognitive systems the way that tools are handled by the motor systems. For this reason, it is important for the brain to take sufficient time to settle on its best interpretation of what just happened, rather than simply its initial, unfinished interpretation. Its carefully refined picture of what just happened is the only thing it will have to work with later.

In conclusion, although many neural models in the literature intimate that a well-defined input is neatly mapped onto a particular output as soon as the leading edge of the information reaches an endpoint (e.g., Zeki & Bartels 1998; Patel et al. 2000; Thorpe et al. 2001), the framework presented here highlights the fact that neural dynamics are influenced through time by the ongoing input of sensory information. We hope this starting point will help navigate us to a physiological explanation of visual awareness.

References

Bachmann, T. (1994). *Psychophysiology of Visual Masking*. Commack, NY: Nova Science.

Blake, R., & Lee, S. H. (2005). The role of temporal structure in human vision. *Behav Cogn Neurosci Rev* **4**(1): 21–42.

Burr, D. C., & Morgan, M. J. (1997). Motion deblurring in human vision. *Proc Biol Sci* **264**(1380): 431–436.

Callaway, E. M. (1998). Local circuits in primary visual cortex of the macaque monkey. *Annu Rev Neurosci* **21**: 47–74.

Crick, F., & Koch, C. (1990). Some reflections on visual awareness. *Cold Spring Harb Symp Quant Biol* **55**: 953–962.

Dennett, D. C. (1991). *Consciousness Explained.* Boston: Little, Brown & Co.

Di Lollo, V. (1980). Temporal integration in visual memory. *J Exp Psychol Gen* **109**(1): 75–97.

Eagleman, D. M. (2001). Visual illusions and neurobiology. *Nat Rev Neurosci* **2**: 920–926.

Eagleman, D. M. (2008). Human time perception and its illusions. *Curr Opin Neurobiol.* In press.

Eagleman, D. M., & Dennett, D. (In preparation). Computation on a need-to-know basis.

Eagleman, D. M., & Sejnowski, T. J. (2000a). Motion integration and postdiction in visual awareness. *Science* **287**: 2036–2038.

Eagleman, D. M., & Sejnowski, T. J. (2000b). Latency difference versus postdiction: response to Patel et al. *Science* **290**(5494): 1051a.

Eagleman, D. M., & Sejnowski, T. J. (2000c). The position of moving objects: response to Krekelberg et al. *Science* **289**: 1107a.

Eagleman, D. M., & Sejnowski, T. J. (2003). The line-motion illusion can be reversed by motion signals after the line disappears. *Perception* **32**: 963–968.

Eagleman, D. M., & Sejnowski, T. J. (2007). Motion signals bias position judgments: a unified explanation for the flash-lag, flash-drag, flash-jump and Frohlich effects. *J Vision* **7**(4): 1–12.

Engel, A. K., Konig, P., et al. (1992). Temporal coding in the visual cortex: new vistas on integration in the nervous system. *Trends Neurosci* **15**(6): 218–226.

Gawne, T. J., Kjaer, T. W., et al. (1996). Latency: another potential code for feature binding in striate cortex. *J Neurophysiol* **76**(2): 1356–1360.

Goodale, M. A., & Milner, A. (1992). Separate visual pathways for perception and action. *Trends Neurosci* **15**: 20–25.

Goodale, M. A., & Milner, A. D. (2004). *Sight Unseen*. Oxford: Oxford University Press.

Hirsh, I. J., & Sherrick, C. E. (1961). Perceived order in different sense modalities. *J Exp Psychol* **62**: 423–432.

Intraub, H. (1985). Visual dissociation: an illusory conjunction of pictures and forms. *J Exp Psychol: Hum Percept Perform* **11**: 431–442.

Julesz, B., & White, B. (1969). Short term visual memory and the Pulfrich phenomenon. *Nature* **222**(194): 639–641.

Kinsbourne, M. (1993). Integrated cortical field model of consciousness. *Ciba Found Symp* **174**: 43–50.

Kline, K. A., & Eagleman, D. M. (2008). Evidence against the snapshot hypothesis of illusory motion reversal. *J Vis* **8**(4): **13**, 1–5.

Kline, K. A., Holcombe, A. O., et al. (2004). Illusory reversal of motion is caused by rivalry, not by perceptual snapshots of the visual scene. *Vision Res* **44**(23): 2653–2658.

Kolers, P., & von Grunau, M. (1976). Shape and color in apparent motion. *Vision Res* **16**: 329–335.

Kopinska, A., & Harris, L. R. (2004). Simultaneity constancy. *Perception* **33**: 1049–1060.

Kopinska, A., Harris, L. R., & Lee, I. (2003). Comparing central and peripheral events: compensating for neural processing delays. *J Vis* **3**: 751a.

Lamme, V. A. (2003). Why visual attention and awareness are different, *Trends Cogn Sci* **7**: 12–18.

Lamme, V. A., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing, *Trends Neurosci* **23**: 571–579.

Libet, B., Alberts, W. W., Wright, E. W., & Feinstein, B. (1967). Responses of human somatosensory cortex to stimuli below threshold for conscious sensation. *Science* **158**: 1597–1600.

Macknik, S. L., & Livingstone, M. S. (1998). Neuronal correlates of visibility and invisibility in the primate visual system. *Nat Neurosci* **1**(2): 144–149.

Maunsell, J. H., Ghose, G. M., et al. (1999). Visual response latencies of magnocellular and parvocellular LGN neurons in macaque monkeys. *Vis Neurosci* **16**(1): 1–14.

Mumford, D. (1994). Neuronal architectures for pattern-theoretic problems. In: C. Koch & J. L. Davis (eds.), *Large-Scale Neuronal Theories of the Brain* (125–152). Cambridge, MA: MIT Press.

Nowak, L. G., & Bullier, J. (1997). The timing of information transfer in the visual system. In: J. Kass, K. Rockland, & A. Peters (eds.), *'Extrastriate Cortex' Cerebral Cortex*, Vol. 12 (205–241). New York: Plenum Press.

Oram, M. W., Xiao, D., et al. (2002). The temporal resolution of neural codes: does response latency have a unique role? *Philos Trans R Soc Lond B Biol Sci* **357**(1424): 987–1001.

Patel, S. S., Ogmen, H., et al. (2000). Flash-lag effect: differential latency, not postdiction. *Science* **290**(5494): 1051.

Pessoa, L., Thompson, E., et al. (1998). Finding out about filling-in: a guide to perceptual completion for visual science and the philosophy of perception. *Behav Brain Sci* **21**(6): 723–748; discussion 748–802.

Pulfrich, C. (1922). Die Stereoskopie im Dienste der isochromen und heterochromen Photometrie. *Die Naturwissenschafte* **10**.

Purpura, K., Tranchina, D., et al. (1990). Light adaptation in the primate retina: analysis of changes in gain and dynamics of monkey retinal ganglion cells. *Vis Neurosci* **4**(1): 75–93.

Purushothaman, G., Patel, S. S., et al. (1998). Moving ahead through differential visual latency. *Nature* **396**(6710): 424.

Ross, J., & Hogben, J. H. (1974). Short-term memory in stereopsis. *Vision Res* **14**(11): 1195–1201.

Schmolesky, M. T., Wang, Y., Hanes, D. P., Thompson, K. G., Leutgeb, S., Schall, J. D., et al. (1998). Signal timing across the macaque visual system. *J Neurophysiol* **79**(6): 3272–3278.

Steinmetz, R., & Engler, C. (1993). Human perception of Media Synchronization. *IBM ENB Tech Rep* **43**(9310). Heidelberg.

Thorpe, S., Delorme, A., et al. (2001). Spike-based strategies for rapid processing. *Neural Network* **14**(6–7): 715–725.

Uttal, W. R. (1979). Do central nonlinearities exist? *Behav Brain Sci* **2**: 286.

VanRullen, R., & Koch, C. (2003). Is perception discrete or continuous? *Trends Cogn Sci* **7**(5): 207–213.

Varela, F. J., Toro, A., et al. (1981). Perceptual framing and cortical alpha rhythm. *Neuropsychologia* **19**(5): 675–686.

Whitney, D., & Cavanagh, P. (2000). Latency difference, not postdiction. *Science*.

Whitney, D., & Murakami, I. (1998). Latency difference, not spatial extrapolation. *Nat Neurosci* **1**(8): 656–657.

Wilson, J. A., & Anstis, S. M. (1969). Visual delay as a function of luminance. *Am J Psychol* **82**(3): 350–358.

Zeki, S., & Bartels, A. (1998a). The asynchrony of consciousness. *Proc R Soc Lond B Biol Sci* **265**(1405): 1583–1585.

Zeki, S., & Bartels, A. (1998b). The autonomy of the visual systems and the modularity of conscious vision. *Philos Trans R Soc Lond B Biol Sci* **353**(1377): 1911–1914.

15

Mechanisms of simultaneity constancy

LAURENCE HARRIS, VANESSA HARRAR, PHILIP JAEKL,
AND AGNIESZKA KOPINSKA

Summary

There is a delay before sensory information arising from a given event reaches the central nervous system. This delay may be different for information carried by different senses. It will also vary depending on how far the event is from the observer and stimulus properties such as intensity. However, it seems that at least some of these processing time differences can be compensated for by a mechanism that resynchronizes asynchronous signals and enables us to perceive simultaneity correctly. This chapter explores how effectively simultaneity constancy can be achieved, both intramodally within the visual and tactile systems and cross-modally between combinations of auditory, visual, and tactile stimuli. We propose and provide support for a three-stage model of simultaneity constancy in which (1) signals within temporal and spatial windows are identified as corresponding to a single event, (2) a crude resynchronization is applied based on simple rules corresponding to the average processing speed differences between the individual sensory systems, and (3) fine-tuning adjustments are applied based on previous experience with particular combinations of stimuli.

15.1 Introduction

Although time is essential for the perception of the outside world, there is no energy that carries duration information, and consequently there can be no sensory system for time. Time needs to be constructed by the brain, and because this process itself takes time, it follows that the perception of when an event occurs must necessarily lag behind the occurrence of the event itself. In fact, Libet (2004) purports to have been able to measure this delay and sets it at about half a second. Several theories have been proposed as to how the perceived timing of events might be reconstructed during this gap, but it is not the aim of this chapter to review these often largely philosophical theories. Here we address a specific aspect of the reconstruction process; how stimuli are judged as being simultaneous. Although the connection between the perception of the absolute time of events ("the now") and actual time is arbitrary, the relative timing of various stimuli is not. Are we able to correct for errors that arise from the fact that various stimuli (both within and across sensory modalities) take different amounts of time to be processed and therefore to be

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

perceived accurately as simultaneous? This chapter reviews the circumstances under which simultaneity is correctly perceived and proposes a model for achieving accurate perception of simultaneity. Although the chemical senses have been demonstrated as having exquisite temporal sensitivity (e.g., differences of 100 msec between the time of arrival of odors to the two nostrils can be detected, von Békésy 1967), this chapter will concern itself specifically with the senses of touch, vision, and hearing.

Several challenges must be overcome to determine which, of all possible stimuli, correspond to a given event and thus are likely to have been simultaneous. The different senses collect data concerning an object or event in different temporal and spatial reference frames, generally about different attributes, and with different degrees of resolution and reliability. Sometimes information picked up by different senses is redundant – for example, information about the time of a handclap can be relayed by auditory, visual, tactile, and proprioceptive systems – which raises the issue of how a single data point (time of clap) might be extracted from such a diverse range of signals.

Determining whether different sensory events occurred at the same time requires the perceptual systems to take into account the different processing times of each sensory signal. Timing differences can arise from both intrinsic and extrinsic factors. Extrinsic factors include the time it takes the energy from the event to reach the neural sensors. Intrinsic factors include the time for the transduction process (King & Palmer 1985) and the neural transmission time for the information to pass from the transducers to the central nervous system (von Békésy 1963; Macefield et al. 1989). Variations in the stimuli themselves, such as their intensity and eccentricity (Wilson & Anstis 1969; Nickalls 1996), also contribute to the variability in processing times that need to be taken into account. Further, personal differences, for example in attention, can change the processing speeds of stimuli (Spence et al. 2001). Considering these factors, it is therefore easy to see how the neural correlates of any two initially simultaneous stimuli can be asynchronous even when the stimuli relate to the same event.

Reconstructing the actual time of an event or the relative timing of its components involves making some allowance for these variable delays in order to identify the simultaneous multimodal components that relate to a given event. We and others have demonstrated that in some situations these variations can indeed be taken into account and stimuli from various modalities can be veridically perceived as synchronous (Engel & Dougherty 1971; Sugita & Suzuki 2003; Kopinska & Harris 2004; Alais & Carlile 2005). This chapter explores how effectively these time variations are dealt with, both intermodally and intramodally, and suggests a mechanism for how this might be done. When the mechanism operates perfectly, the relative timing of sensory events is correctly perceived.

15.2 Simultaneity constancy

The ability to perceive simultaneous events correctly despite variations in the timing of the sensory representations of the component stimuli is known as simultaneity constancy (Kopinska & Harris 2004). Simultaneity constancy is in line with other perceptual constancies, for example, size constancy in which the perceived size of an object is maintained despite variations in the retinal information (Gregory 1963; McKee & Smallman 1998; see Walsh & Kulikowski 1998 for a review).

15.3 Assessing the challenge

Before we consider how well incidental temporal delays between sensory signals are compensated by the brain, we need to assess the magnitude of the temporal asynchronies that need to be compensated. Some delays can be objectively measured, such as the time it takes for sound to reach the observer, but delays introduced internally by differential processing times are more difficult to assess. Theoretically, objective measures can be obtained from delays in evoked potentials to unimodal stimuli (Jeffreys & Axford 1972; Celesia 1976; Lesevre 1982; Liegeois-Chauvel et al. 1991). But signals arrive at different parts of the cortex at a range of times, so the measurable response might not correspond to the signals involved in temporal judgments.

Reaction times to stimuli can provide an indirect measure of processing time. When used with care they provide a relatively straightforward method for assessing the effects of stimulus characteristics on neural processing time (Luce 1986). Reaction times are composed of several components: the time it takes energy to reach the receptors, the transduction and conduction times, the central processing time, and the time it takes to plan and execute the motor response. There is no logical reason to think that the motor component would be affected by manipulation of stimulus properties such as intensity. Therefore, we assume that the difference in reaction times to various individual stimuli represents the total time difference that needs to be compensated for to resynchronize stimuli. A reaction time difference is an estimate of the staggered delay with which the stimuli arrive and meet the detection criterion; a simultaneity constancy mechanism needs to correct for this delay because it was not present in the originally synchronous stimuli.

Similar to evoke potentials, however, the reaction time for detecting simple stimuli may not reflect the processing times used for tasks other than detection (e.g., Jaśkowski 1999; Jaśkowski & Verleger 2000). Further dissociations between reaction time and temporal order judgments (TOJs) can occur if the neural responses to the stimuli are of different magnitudes. Because reaction times rely on a stimulus meeting a criterion before a response is made, it is important that the magnitude of the internal representations are approximately equal (Sternberg & Knoll 1973): "Unless two responses are identical in size and shape, differing by a time translation only, there is not uniquely defined latency difference" (p. 649). It is therefore important for experiments using this method to equate stimuli for intensity as much as possible.

15.4 Methods used to assess perceived simultaneity

There are two methods generally used to assess the perception of simultaneity among auditory, visual, and tactile stimuli:

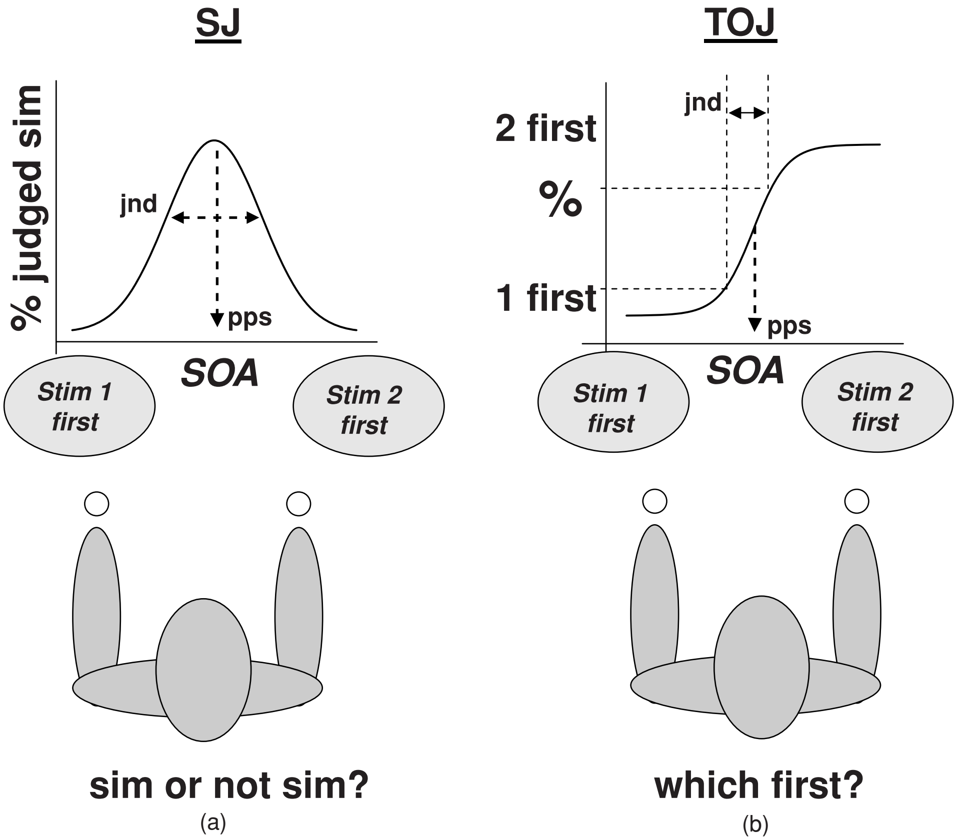
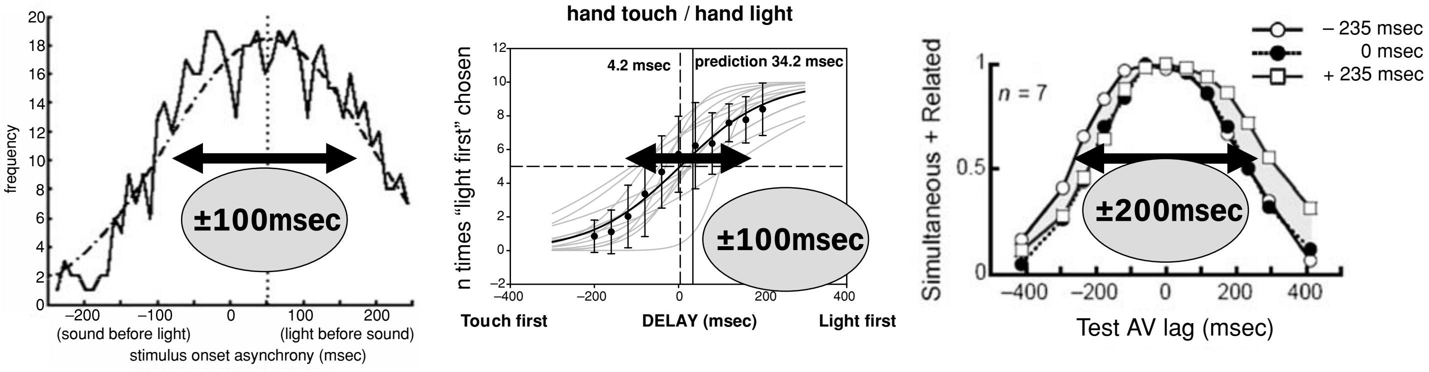


Fig. 15.1 Two methods to assess simultaneity. (a). Simultaneity judgments (SJs). The SJ data graph shows the percentage of times a given SOA was chosen as appearing simultaneously. The curve is well approximated by a Gaussian where the point of subjective simultaneity (PSS) is given by the peak of the curve and the standard deviation (STD ±34%) is equivalent to the just noticeable difference (JND). (b). Temporal order judgments (TOJs). The TOJ data graph shows the percentage of times that stimulus 2 (e.g., a light) was chosen as appearing before stimulus 1 (e.g., a sound). This function is well approximated by a cumulative Gaussian. The PSS is given by the 50% point where the subject is equally likely to say either stimulus came first, and the STD is equivalent to the JND defined as ±34% from 50%. Note that both these functions assume symmetry around the PSS, which may not be a valid assumption under all circumstances.

1. *Simultaneity judgments (SJ):* A forced-choice decision is made between whether two stimuli are "simultaneous" or "successive." Generally these decisions are reported as a frequency distribution of the "number of times subjects reported simultaneous" that tends to be normally distributed when plotted as a function of the stimulus onset asynchrony (SOA) between the two stimuli (Fig. 15.1(a)). The peak of this curve indicates the SOA at which subjects are most likely to say "simultaneous": The point of subjective simultaneity (PSS). It is convenient to define the just noticeable difference (JND) as one standard deviation (+34%) from the PSS. JNDs so defined typically indicate that combinations of auditory, visual, and tactile stimuli need to be separated



(a) *from Stone et al. (2001)* (b) *from Harris and Harrar (2005)* (c) *from Fujisaki et al. (2004)*

Fig. 15.2 The width of the temporal integration window. Several studies using different methods suggest that the window within which temporal integration occurs is between 100 and 200 msec. (a) Auditory visual simultaneity judgments. Data reproduced by permission from Stone, R. V. et al. (2001). "When is now?" Perception and simultaneity. *Proc Roy Soc Lond* B 268: 31–38. (b) Visual tactile temporal order judgments. Data reproduced with permission from Harrar, V, & Harris, L. R. (2005). Simultaneity constancy: detecting events with touch and vision. *Exp Brain Res* 166: 465–473. (c) Auditory visual simultaneity judgments before and after exposure to an adaptation regime designed to shift the PSS. Note the asymmetric changes from the baseline (filled circles) with the dominant effect on JNDs. Data reproduced with permission from Fujisaki et al. (2004). Recalibration of audiovisual simultaneity. *Nat Neurosci* 7: 773–778.

by about 150–200 msec to be reliably perceived as asynchronous (see, e.g. Figs. 15.2(a) and (c), taken from Stone et al. 2001, and Fujisaki et al. 2004, respectively).

In interpreting SJ data, care has to be taken to distinguish between when subjects are asked (a) whether the members of a single stimulus pair were simultaneous or not (e.g., Zampini, Guest, et al. 2005) or (b) which of two successively presented pairs appeared most simultaneously. The former is highly subjective, psychophysically uncontrolled, and subject to criterion shifts in the JND. The latter design requires that one of the sets be "truly simultaneous." It therefore cannot be used to determine the PSS because it requires the PSS to be presented as the reference for each forced choice. It can, however, allow one to determine relatively criterion-free estimates of the JNDs.

2. *Temporal order judgments (TOJ):* A forced choice is made as to which of two stimuli came on first. A psychometric curve is fitted to the proportion of times a subject perceived one of the stimuli as occurring first, plotted as a function of the SOA (Fig. 15.1(b)). Such functions can be well described by a cumulative Gaussian with the 50% point defining the PSS and the standard deviation defining the JND. JNDs as such are typically around ±50 msec (Spence et al. 2001; and see Fig. 15.2(b), taken from Harrar & Harris 2005).

Care has to be taken to consider biases – for example, response bias such as a tendency to respond with the right hand, or a bias toward responding "light" as it is generally the most salient cue. Biases can be reduced by alternating blocks in which subjects report which one came first with blocks in which they report which one came second.

15.5 Comparison of SJs and TOJs

SJ and TOJ measurements provide estimates of the PSS and associated JNDs. An advantage of measuring SJs is that direct judgments of perceived simultaneity are obtained that are relatively free of cognitive bias (Schneider & Bavelier 2003; Zampini, Shore et al. 2005). A disadvantage of SJs is that judgments of whether two things appear simultaneously or not are highly subjective and vulnerable to random fluctuations of criterion and bias. The participant's criteria for perceived simultaneity may not be consistent throughout or between experimental sessions. An advantage of TOJs is that they provide statistically reliable psychometric data relatively immune to subjective criteria biases because of the forced choice between two independent alternatives; a disadvantage is that the decision is based on the remembered (i.e., retrospective) temporal sequence and is therefore vulnerable to postperceptual biases.

At first glance it would seem that the PSS and JND obtained by SJs and TOJs would be the same. But in fact the values obtained using SJs and TOJs are not necessarily the same. It is possible to perceive two stimuli as asynchronous but to not know which one came first. Allan (1975) suggested that SJs and TOJs are processed at different stages and that successiveness is needed before correct temporal order can be perceived.

SJs and TOJs can be affected differently by certain properties of the stimulus pairs (e.g., Shore et al. 2002). For example, spatial redundancy: When multimodal stimuli are separated both by modality and space, multimodal TOJs are facilitated but SJs are impeded (Spence et al. 2003). SJs and unimodal TOJs are easiest when stimuli are in close spatial proximity probably because apparent motion can be used to cue temporal order and asynchrony and because detecting apparent motion is easier when the stimuli are close together.

Thus the choice of measure should be carefully tailored to the particular experimental question and stimuli used. Some problems plague both techniques, however. For example, JNDs are always vulnerable to the phenomenon of temporal ventriloquism whether assessed by SJs or TOJs. As a parallel to spatial ventriloquism, in which the perceived location of a sound is altered by the presence of a related light, temporal ventriloquism is when the perceived *time* of a light (or other stimulus) is affected by the presence of a related sound (or other stimulus) (Aschersleben & Bertelson 2003; Bertelson & Aschersleben 2003; Morein-Zamir et al. 2003; Jaekl & Harris 2007). If stimuli are drawn together in time in this way, they will tend to be perceived as simultaneous over a larger range than they might otherwise be – causing the JND to appear larger. Temporal ventriloquism as such only affects the JND. However, if the effects of temporal ventriloquism are asymmetric (Morein-Zamir et al. 2003), then the PSS could also be shifted closer to true simultaneity (see also Spence & Squire 2003). It is therefore not possible to measure the accuracy (JNDs) by which the relative timing of the two components is assessed by the intervention of separating them by a small amount of time because the act of separating them can create a different percept in which the components' perceived location in time is distorted. Only if the factors determining temporal ventriloquism are kept constant can two sets of responses be compared. We will now review some of the experiments using SJs and TOJs to investigate the perceived relative time of various stimulus pairings.

15.6 Simultaneity constancy in the auditory/visual system

Although light travels enormously faster than sound, the complexity of the transduction process for light takes about 45 msec longer than sound (vision about 60 msec: Jeffreys & Axford 1972; Lesevre 1982; auditory about 15 msec: Celesia 1976; Liegeois-Chauvel et al. 1991). The speed of travel of sound adds a delay of 3 msec/m so that at 15 m the faster transduction speed and the slower travel time roughly cancel out and auditory and visual information reach the brain at approximately the same time. This theoretical concept has been called "the horizon of simultaneity" by philosophers such as Pöppel (1988) and Dennett (1991). When TOJs for sound/light pairs at different distances were made to measure this "horizon" it was interesting to notice that there was, in fact, nothing special about auditory/visual pairs presented at this distance. The relative timing of pairs over a whole range of distances up to 32 m was correctly perceived (Engel & Dougherty 1971; Kopinska & Harris 2004). A simultaneity constancy mechanism thus seems to be active over this range of distances. Using headphones while watching visual targets at different distances, Sugita and Suzuki (2003) confirmed that the time difference needed to match the auditory and visual targets increased with distance up to about 20 m. Their data are reproduced in Fig. 15.3. Because Sugita and Suzuki's subjects wore headphones, the increasing delay of the sound required for it to appear simultaneous with a light at increasing distance corresponds to the brain allowing for the time it would have taken the

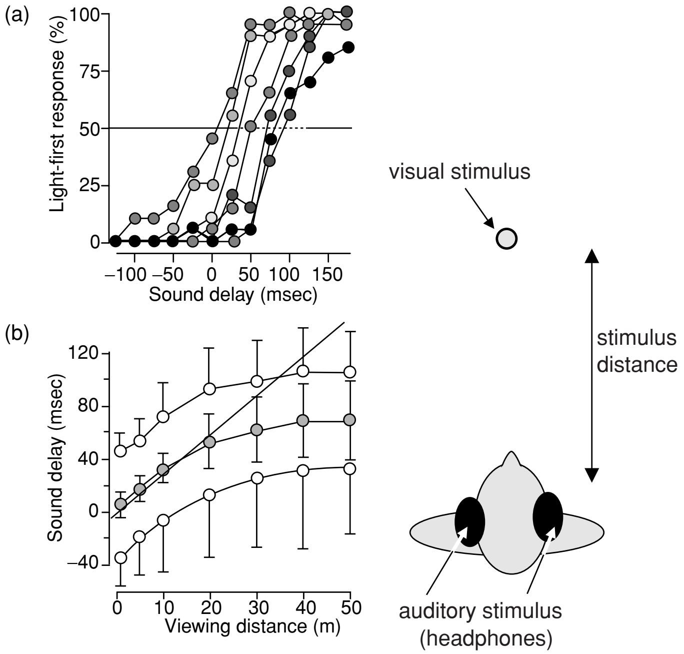


Fig. 15.3 Data taken from Sugita and Suzuki (2003) showing TOJs obtained for sound presented through headphones and lights presented at different distances (top panel). The lower panel shows the PSS obtained from each curve plotted as a function of distance (filled circles). There is an increase in PSS with distance up to about 15 m even though the sounds were presented in headphones. This increase corresponds to the brain allowing for the time it would have taken sound to reach the head from the light source if the sound had been externalized. Reproduced with permission from Sugita and Suzuki (2003). Audiovisual perception: implicit estimation of sound-arrival time. *Nature* 421: 911.

sound to reach the head from the location of the light. This is a powerful demonstration of simultaneity constancy driven by the visual cues to distance.

Kopinska and Harris (2004) found that many factors that introduce differences in the timing of the light and sound stimuli (eccentricity, distance, and intensity) were all taken into account up to 32 m so that the simultaneity of many stimulus pairs was correctly perceived up to this distance (Fig. 15.4).

Under circumstances where distance information is not easily available, such as in a dark anechoic chamber (Lewald & Guski 2004) or when causality is involved (Arnold et al. 2005), simultaneity constancy is not always found. Lewald and Guski's data are reproduced in Fig. 15.5. An example involving causality is when two disks move on a collision course; they can either be seen to pass through each other or to bounce off each other (Arnold et al. 2005). The bounce interpretation is perceived more often if a sound accompanies the "collision." The sound is most effective in doing this if it is presented not at the instant of collision but just prior to it – suggesting that the sound and the visual event need to match in "brain time" rather than in real time for optimum performance. Interestingly, Arnold

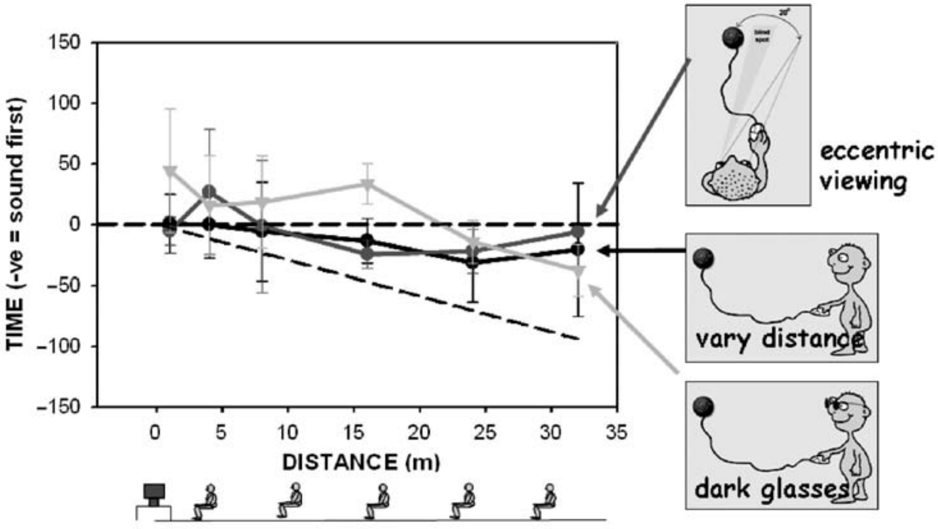


Fig. 15.4 Data redrawn from Kopinska and Harris (2004) showing the PSSs obtained from TOJs between lights and sounds presented at distances of up to 32 m. The PSSs are not significantly different from zero and are not significantly shifted by manipulations that dramatically alter the speed with which light is processed (eccentric viewing or wearing dark glasses).

et al. (2005) found that the most effective stimulus onset asynchrony (SOA) for perceiving the bounce needed to be increased with the distance to the observer in accordance with the longer time taken for the sound to reach the observer. Thus it appears that the simultaneity mechanism is not engaged in circumstances involving the perception of causality.

Alais and Carlile (2005) showed that robust cues to distance are needed for auditory/visual resynchronization. Of course, if distance cues are not available, such as in an anechoic chamber, or in a field, then it would be next to impossible to correctly compensate for the delayed auditory stimulus. Dixon and Spitz (1980) determined the amount of temporal delay for which the soundtrack of a movie could be desynchronized and still be perceived as synchronous. Their results showed a large tolerance within which the desynchronization went unnoticed. However distance cues in movies, so important for simultaneity constancy to be achieved, almost always represent a conflict because the distance from the viewer to the screen is rarely the distance at which the filmmaker wishes the audience to perceive the action being portrayed.

15.7 Simultaneity constancy in the visual/tactile system

Sensory timing differences also need to be taken into account for the veridical perception of simultaneous visual and tactile stimuli. Tactile stimuli, like auditory stimuli, are transduced

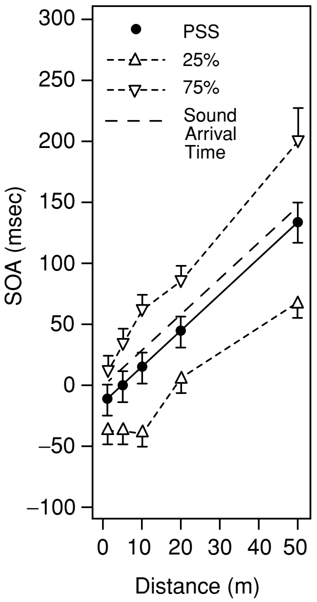


Fig. 15.5 Data taken from Lewald and Guski (2004) showing that, when distance is hard to determine (in this case in a dark anechoic room), PSS increases with distance (for lights and sounds presented at different distances) and does not show simultaneity constancy. Note that, in this example, when the stimuli are presented at a distance of zero, the sound needs to be turned on slightly before the light to be heard as simultaneous, rather than 45 msec after to allow for its faster processing time, suggesting that some compensation has taken place. Reproduced with permission from Leward, J., & Guski, R. (2003). Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. *Cogn Brain Res* 16: 468–478.

faster than visual stimuli (King & Palmer 1985; Pöppel et al. 1990). For sounds, we noted that the longer time of transmission from the source to the end organ tended to offset this time advantage. For touches, the distance from the point of tactile stimulation to the brain determines the delay, which depends on the length of the nerves to the various parts of the body (Macefield et al. 1989).

To determine if simultaneity constancy was applied to visual and tactile timing differences, Harrar and Harris (2005) measured TOJs between lights and touch stimulation on the hand. This point of tactile contact was chosen because stimulation on the hand is processed approximately 34 msec faster than visual stimuli, as measured by reaction time differences. The PSS values for lights and touches on the hand were not significantly different from zero, implying that simultaneity constancy was achieved.

When a touch and a visual stimulus are synchronous but occur on different parts of the body they are more likely to have separate causes or correspond to separate events. Indeed, spatial congruency may be required, and in some cases has been found to be essential, for stimuli to be bound into a single event and simultaneity constancy activated (Spence

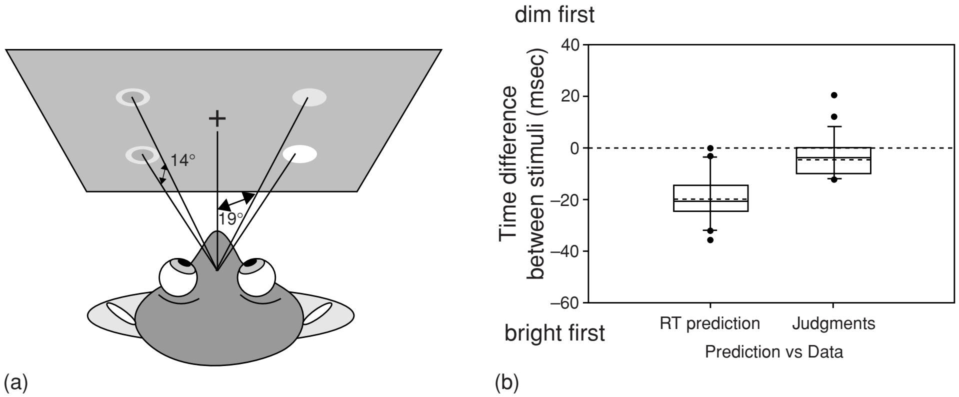


Fig. 15.6 Subjects were shown dim and bright lights (a) and asked which one came first. Reaction times to the dim lights were about 20 msec than for the bright ones as shown in the bar chart on the right (b) but the PSSs obtained from TOJs were not significantly different from zero.

et al. 2001, 2003; Spence & Squire 2003; Soto-Faraco et al. 2004; Harrar & Harris 2005; Zampini, Brown et al. 2005; Zampini, Guest et al. 2005).

15.8 Simultaneity constancy within a modality

The simultaneity-detecting mechanism does not only need to deal with timing differences arising between senses but also within a sense. A single event usually has different attributes, for example, color, movement, and shape, that are processed with different speeds even by a single sense, for example, vision.

Therefore there is also a need for an intramodal simultaneity constancy mechanism within vision. In the experiment illustrated in Fig. 15.6, Kopinska and Harris used targets with different luminances. Reaction times suggested that these patches were processed at different speeds, but TOJs showed that the timing difference was accounted for and simultaneity constancy was achieved, that is, the PSS was not significantly different from true simultaneity. These results provide evidence that simultaneity constancy is achieved between separately processed visual signals and demonstrate that the adjustment of temporal information occurs also within a single modality.

Reaction times to touches are longer for body parts that are further from the center of the head (Bergenheim et al. 1996; Harrar & Harris 2005), and the perceived temporal order of two touches on the body is predictable not from which body part was actually touched first but from the relative distances of the two touches from the head (Craig & Baihua 1990; Bergenheim et al. 1996; Shore et al. 2002; Harrar & Harris 2005). During active touches, when the foot is touched by the hand, for example, rather than having the two sensations at different times, as they would if simultaneity constancy were not engaged, one of the two touches is inhibited (von Békésy 1963). Thus the touch system seems to have found other ways of dealing with timing differences. Bergenheim et al. (1996) suggested that the touch system has "a degree of precision that is functional for the individual."

15.9 Flexibility

Fujisaki et al. (2004) showed that the PSS for a given comparison is not fixed. The PSS of a light/sound pair can be shifted following a period of repeated exposure to a light/sound pair separated by a particular delay (Fujisaki et al. 2004; see also Vroomen et al. 2004). These results may be interpreted as indicating that the system is able to learn a new interstimulus time difference as corresponding to simultaneity. Adaptability of this kind suggests that the mechanism is not hardwired but is influenced by learning.

Fujisaki et al. (2004) adapted subjects to several different temporal intervals and found that responses varied depending on the previously experienced SOA. However, in some conditions (see Fujisaki et al.'s Figs. 15.2(b) and 15.2(d)), subjects' PSSs were all in the direction of "sound first" even after adapting to "light first" stimuli. Further, many shifts (relative to the "no adaptation" condition) were actually in the direction opposite to the actual delay. Similar results were found in Harrar and Harris (2005) (see Fig. 15.5). These results suggest that there may be a bias toward shifting the audio/visual PSS in the direction of sound first, but it is not yet known why. It could be because of attention shifts (Spence et al. 2001) that might only occur when both stimuli are present (Aschersleben 1999) or because of an expectancy (or "prior") built into the system based on the usual relative arrival times of light and sound. Miyazaki et al. (2006) modeled the effect of prior expectancies on the adaptability of the simultaneity constancy mechanism and showed how PSS shifts for tactile stimuli can be predicted statistically either toward or away from an experienced time stagger. Despite these complications, the flexibility of the simultaneity system provides an experimental tool for testing numerous, still outstanding questions about the simultaneity mechanism.

15.10 A single, global mechanism or multiple, specific mechanisms?

There may be a single, global mechanism responsible for simultaneity constancy both within and between modalities, or there may be multiple simultaneity mechanisms comprising separate systems for audio/visual, visual/tactile, and audio/tactile comparisons. These two alternatives can be distinguished by exploiting the flexibility demonstrated by Fujisaki et al. (2004), adapting the PSS of an audio/visual pair and looking for any effects in audio/tactile and visual/tactile pairs.

15.10.1 Implications from PSS shifts

Harrar and Harris (2008) tested TOJs before and after adaptation to each of three combinations of time-staggered stimuli (light/sound, light/touch, sound/touch), to distinguish a single global simultaneity constancy mechanism from multiple mechanisms for different

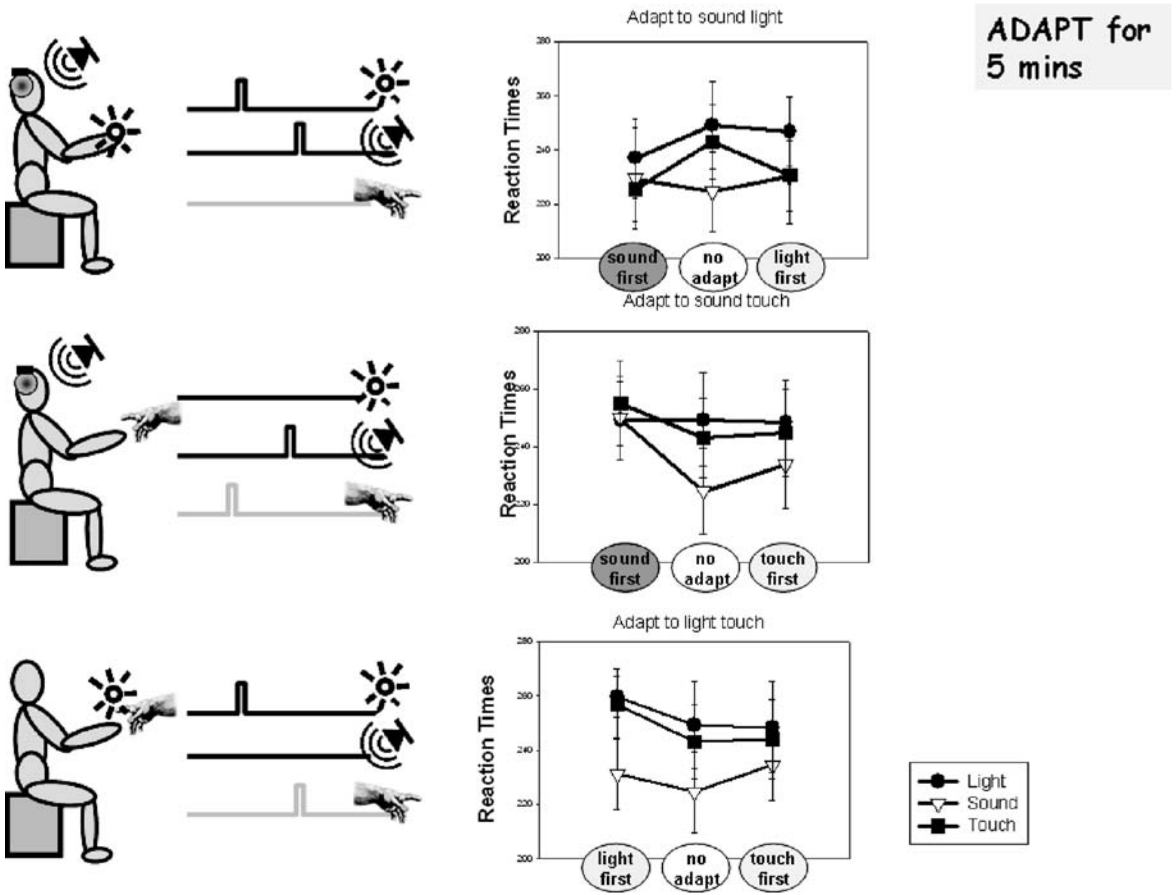


Fig. 15.7 Subjects were exposed to 5 min of time-staggered stimulus pairs (sound/light, sound/touch, or light/touch), as shown in the cartoons on the left. Such exposure to sound/light pairs results in a shift of the PSS (see text). However, no shifts were found in the reaction times to sounds, lights, or touches after any pattern of exposure as shown by the graphs on the right.

stimulus combinations. A PSS shift in a pair that was not the one exposed would support a global mechanism, whereas a PSS shift of only the exposed pair would support multiple separate simultaneity constancy mechanisms. After adapting to a time-staggered sound/light combination, PSSs for the light/sound pair shifted but the PSS for the sound/touch and light/touch pairs did not change (Harrar & Harris 2008). This pattern of results renders a single, global simultaneity constancy mechanism quite unlikely. Therefore we suggest a multichannel system.

Although it appears that intermodal comparisons are accomplished through separate mechanisms, intramodal comparisons may each be executed by a single mechanism. There is some generalization across audiovisual stimuli, for example, from complex (watching a video of a person talking) to simple (beeps and flashes) examples of the same multimodal stimulus pair, suggesting that audiovisual comparisons share a simultaneity mechanism (Navarra et al. 2005; Vatakis & Spence 2006).

15.10.2 Implications from reaction times

Harrar and Harris (2008) also, for the first time, tested reaction times to individual, unimodal stimuli before and after adaptation to each of three combinations of time-staggered stimuli. Any reaction time change would indicate a global, low-level change in the functioning of the system that would blindly and without regard to context operate with a particular delay.

Even though the PSS of a light/sound pair shifted after repeated exposure, there were no systematic changes in reaction times to any of the individual stimuli (Fig. 15.7; Harrar & Harris 2008). This suggests that the timing corresponding to PSSs is constructed centrally and not tied to timing differences in the sensory signals. This is consistent with the simultaneity constancy mechanisms not working on the individual stimulus level.

15.11 Is touch special?

When Harrar and Harris (2008) tested TOJs for each of three combinations of time-staggered stimuli (light/sound, light/touch, sound/touch), the PSS of the light/touch pair or the sound/touch pair never changed, even after asynchronous exposure to light/touch or sound/touch pairs. In other words, the perception of the relative timing of stimuli where touch is involved seems more rigid. The touch system may be fundamentally different from the more passive auditory and visual systems because of its association with actions and knowledge of limb position, or it might be related to the fact that the temporal properties of a touch do not need to take into account outside factors (Miyazaki et al. 2006). The lack of plasticity in the tactile system may be evidence for a fixed temporal somatosensory homunculus (Bergenheim et al. 1996). Alternatively, the lack of immediate plasticity within the touch system may be related to the inhibition of one touch by another as described by von Békésy (1963).

15.12 Mechanisms of simultaneity constancy

Given that simultaneity constancy is engaged under many circumstances, it now behooves us to explain how variations in timing might be allowed for in a flexible simultaneity constancy mechanism. There are two broad classes of models that could achieve simultaneity constancy. Signals could be brought into line at some point in the brain using computations based on knowledge of the various factors that might desynchronize the signals. An alternative model is based on probability. Windows of acceptance for a range of time differences between stimulus pairs are generated based on previous experience. In the probabilistic model, if the time difference between signals falls in the acceptance window, the two signals are regarded as simultaneous. Both these classes of models can be implemented with single or multiple comparators.

15.12.1 Computational models

Computational models for resynchronizing individual stimuli have a long history that often smacks of the Cartesian theater in which an inner homunculus experiences what the senses present in some kind of "show in the brain" (see Dennett 1991 for a comprehensive debunking of this idea). Sternberg and Knoll (1973) described several variations. In these models, stimuli arrive at the brain at variable times. However, before they arrive at the decision center, the individual stimuli have their delays "corrected" (see Sternberg & Knoll's Fig. 12(a)). Computational models require altering the timing of information as it passes through the nervous system such that by the time it arrives at the relevant decision-making site, any timing differences have already been removed.

If, for example, to achieve a central alignment in the representation of stimuli in the brain, the processing time of a particular modality needed to be made longer, this added delay in processing time would subsequently be applied in all circumstances in which that stimulation occurred regardless of which other stimuli were present. For example, if the processing of light were delayed to compensate for the time it takes for a sound coming from a distant event, the added delay would affect all subsequent perceptions involving visual stimuli. Delaying the processing of visual stimuli, while having the advantage of bringing it into synchrony with some particular sounds, would generally seem to be risky. Alternatively, computations could be simplified depending on which stimuli are being compared, adding different delays for touch if it is being compared with a light or a sound (see Sternberg & Knoll's Fig. 12(b,c)) and limiting the operational range (Dennett & Kinsbourne 1992).

Reaction time and processing time evoke potential differences between various stimuli and suggest empirically that no peripheral compensation for timing differences takes place at the individual stimulus location; however, this might be due to the dissociation between reaction times and perceived time. Intuitively it seems unlikely that all the sources of variation in timing could be hardwired into the system or that it would be advantageous for it to be so. We therefore need to consider another class of model.

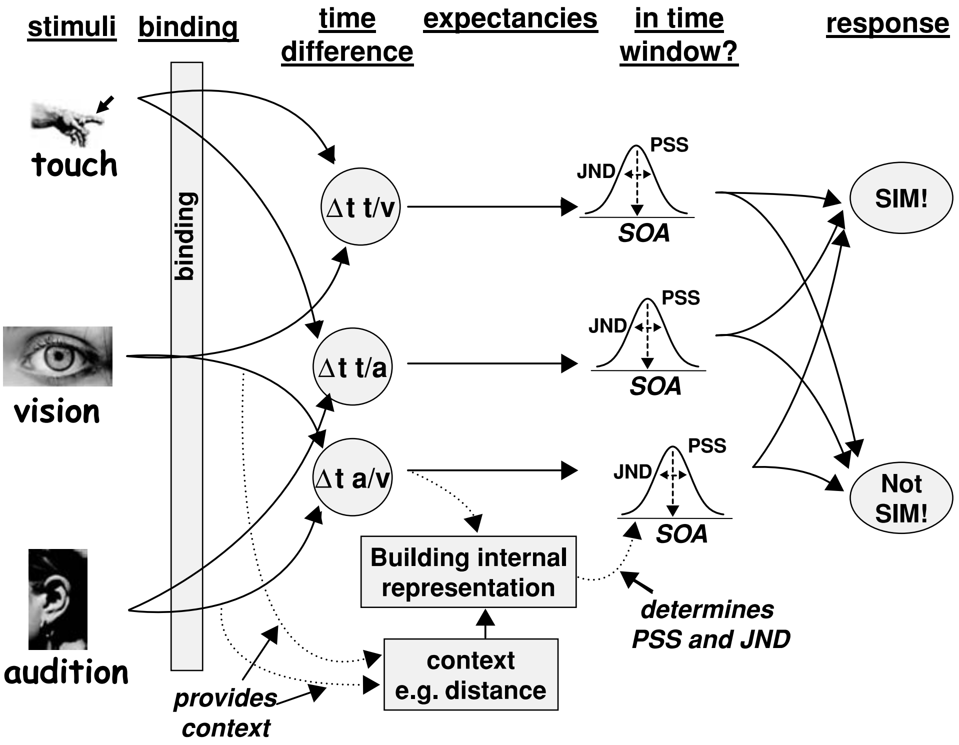


Fig. 15.8 Model for achieving simultaneity constancy. Information from different senses about which temporal judgments are to be made are first "bound," meaning that they are regarded as coming from the same event. The time difference between them is noted and used for two purposes: (1) to decide if the difference falls within the range regarded as simultaneous (rightward pointing arrow), and (2) to contribute to the probabilistic assessment of what that range should be (down right pointing dotted arrow). For each stimulus combination an "internal representation" is built and available for comparison, dependent on experience and context. These processes are only drawn for audiovisual comparisons but are probably essentially the same for touch–vision and touch–audition, although the "contributing to the assessment" arrow from the tactile system may be weak or nonexistent. The input from the time-difference assessment represents an instantaneous updating of this internal representation.

15.12.2 Probability models

Our suggested model for achieving simultaneity constancy is probabilistic rather than computational. Figure 15.8 summarizes the basic steps. The first part addresses the question of selecting appropriate stimuli for temporal compensation. The stimuli about which judgments are to be made are selected from the entire array of stimuli by a binding process. Although it is yet unclear how corresponding stimuli come to be bound together, there seem to be some spatial and temporal criteria (Spence et al. 2001; Bertelson & Aschersleben 2003).

The next stage in the model notes the timing difference between bound stimuli and uses this for two purposes. First, the time difference is used to generate an internal expectation of the time differences associated with those stimuli tagged for the particular context in which the stimuli occurred, for example, at a particular distance. It is the effect of context on this expectation that allows the "constancy" aspect of the model. Expectancies are continuously updated based on the context so that veridical and constant simultaneity can be perceived despite changes in the "context." Note that the polarity of the time difference needs to be kept as part of this tagging procedure. Just as when describing experimental results, it is important to define that, for example, positive corresponds to "sound first." Second, the time difference is also compared to this internal representation of the time differences associated with the stimuli in that particular context (shown in Fig. 15.8). If the time difference of the signals in question falls within the relevant time window, a decision of "simultaneous" is made. The time window determines the probability with which the stimuli are perceived as simultaneous. Because there is evidence for multiple simultaneity mechanisms (see *Implications from PSS shifts* previously), these separate mechanisms are distinguished within our model (Fig. 15.8).

To work through an example, then, the processing time of touches on the body is generally faster than that of lights. So when a multimodal stimulus occurs that involves tactile and visual stimulation, such as watching something touch the skin or looking at an object being manipulated in the hand, the tactile input that leads the visual input, by roughly 40 msec, will be compared with the probability function (which has a peak at 40 msec based on previous experience of this combination) and will thus be within the simultaneity window. Lights and touches with a delay of 40 msec will thus be identified as simultaneous.

15.12.3 Recalibration achieved by JND increase followed by PSS shift

How does our model become "recalibrated" during exposure to a new time stagger introduced experimentally as demonstrated by Fujisaki et al. (2004)? Navarra et al. (2005, 2007) were able to find an adaptation effect for both audio/tactile and visuo/tactile pairs, even though their pairs included touch. The JNDs increased after adaptation, thus tending to include a larger range of times (potentially including the experienced time stagger) and causing more SOAs to be perceived as "simultaneous." Navarra et al. (2005) proposed that expanding the JND is a precursor for all adaptive changes of the PSS. Figure 15.9 shows the internal expectation for time delays ("ΔT old" in Fig. 15.9: The sum of previous experience – dotted curve) changes when repeatedly exposed to a new specific delay (labeled "ΔT new" in Fig. 15.9). It assumes that the stored expectation declines at some rate and can be replaced as new data accumulate about a new probability of what time differences correspond to simultaneous. An interesting feature is that, during the transition phase, the PSS shift is, initially, the result of an asymmetric increase in JND (simply caused by adding the distributions of the initial and the new representation). Such an asymmetric expansion is visible in the data of Fig. 15.2(c) (taken from Fujisaki et al. 2004 during an adaptation regime). Over time, and with further experience, the model predicts that the JNDs get

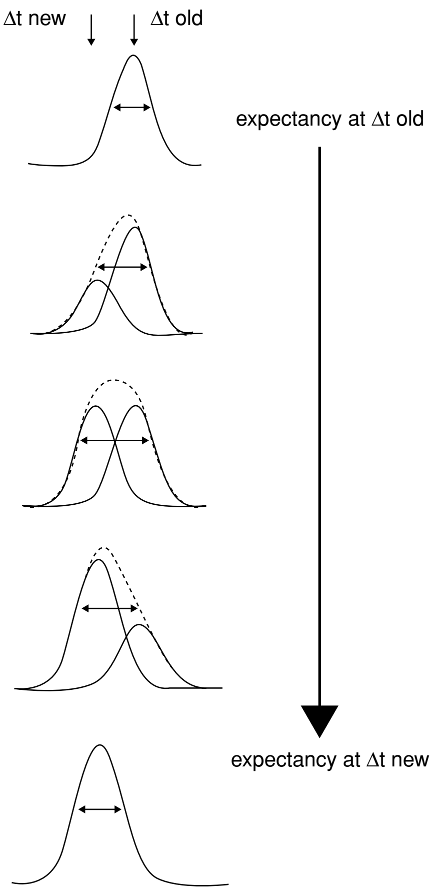


Fig. 15.9 Building an internal representation: changing one expectation to another. The curve at the top represents the expectation that Δt old corresponds to simultaneity. After Δt new is repeatedly experienced, the expectancy changes to this new value. In between, the window becomes wider and asymmetric before settling down to the new values.

smaller again, but the largest PSS shift (as compared with the PSS shifts during the adaptation) still remains. This model therefore explains why similar experiments find different results with regard to JND changes and PSS shifts. PSS shifts found without JND changes (Fujisaki et al. 2004; Vroomen et al. 2004) signify the end of the recalibration process, whereas small PSS shifts accompanied by large JND increases (Navarra et al. 2005, 2007) suggest an early stage of the recalibration process.

15.13 Conclusions

This chapter has reviewed the occurrence of an important perceptual phenomenon: simultaneity constancy. There appear to be several parallel simultaneity constancy mechanisms, each working on the timing of different combinations of stimuli. We have developed a model to explain how this may be done. The rules of when to activate these systems are not clear, but binding within a spatial window seems to be required. In this model, when a simultaneity mechanism is activated, the timing difference is compared with an internal representation of the expected time difference for those particular stimuli. SJ and TOJ estimates can then be made about the relative timing of stimuli. The model allows simple contextual factors, such as distance, to affect the internal expectancy and thus the perceived relative time of the stimuli. Highly cognitive tasks, such as watching a ball bounce, do not seem to engage simultaneity constancy, whereas simple beeps, flashes, and videos do. This may have to do with the level of complexity of the context within which the stimuli occur or might reflect the operation of other principles about which we are currently ignorant. Understanding the brain's abilities and limitations and the way in which some of the principles outlined here could be neurally implemented will advance our understanding of the perception of multimodal time.

15.14 Acknowledgments

This work was supported by a grant from the Natural Science and Engineering Research Council of Canada to Laurence Harris. Vanessa Harrar and Philip Jaekl hold NSERC graduate scholarships. We are grateful to Sugirhini and Shamini Selvanayagarajah for their help with the data collection.

References

Alais, D., & Carlile, S. (2005). Synchronizing to real events: subjective audiovisual alignment scales with perceived auditory depth and speed of sound. *Proc Natl Acad Sci U S A* **102**: 2244–2247.

Allan, L. G. (1975). The relationship between judgments of successiveness and judgments of order. *Perception & Psychophysics* **18**: 29–36.

Arnold, D. H., Johnston, A., & Nishida, S. (2005). Timing sight and sound. *Vision Res* **45**: 1275–1284.

Aschersleben, G. (1999). Task-dependent timing of perceptual events. In G. Aschersleben, T. Bachmann, & J. Müsseler (eds.), *Cognitive Contributions to the Perception of Spatial and Temporal Events* (293–318). North Holland: Elsevier.

Aschersleben, G., & Bertelson, P. (2003). Temporal ventriloquism: cross-modal interaction on the time dimension. 2. Evidence from sensorimotor synchronization. *Int J Psychophysiol* **50**: 157–163.

Bergenheim, M., Johansson, H., Granlund, B., & Pedersen, J. (1996). Experimental evidence for a sensory synchronization of sensory information to conscious experience. In S. R. Hameroff, A. W. Kaszniak, & A. C. Scott (eds.), *Towards a Science of Consciousness: The First Tucson Discussions and Debates* (301–310). Cambridge, MA: MIT Press.

Bertelson, P., & Aschersleben, G. (2003). Temporal ventriloquism: crossmodal interaction on the time dimension. 1. Evidence from auditory-visual temporal order judgment. *Int J Psychophysiol* **50**: 147–155.

Celesia, G. G. (1976). Organization of auditory cortical areas in man. *Brain* **99**: 403–414.

Craig, J. C., & Baihua, X. (1990). Temporal order and tactile patterns. *Perception & Psychophysics* **47**: 22–34.

Dennett, D. C. (1991). *Consciousness Explained.* Boston: Little, Brown and Company.

Dennett, D. C., & Kinsbourne, M. (1992). Time and the observer: the where and when of consciousness in the brain. *Behav Brain Sci* **15**: 183–201.

Dixon, N. F., & Spitz, L. (1980). The detection of auditory visual desynchrony. *Perception* **9**: 719–721.

Engel, G. R., & Dougherty, W. G. (1971). Visual-auditory distance constancy. *Nature* **234**: 308.

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nat Neurosci* **7**: 773–778.

Gregory, R. L. (1963). Distortion of visual space as inappropriate constancy scaling. *Nature* **199**: 678–680.

Harrar, V., & Harris, L. R. (2005). Simultaneity constancy: detecting events with touch and vision. *Exp Brain Res* **166**: 465–473.

Harrar, V., & Harris, L. R. (2008). The effect of exposure to asynchronous audio, visual, and tactile stimulus combinations on the perception of simultaneity. *Exp Brain Res* **186**: 517–524.

Jaekl, P. M., & Harris, L. R. (2007). Auditory-visual temporal integration measured by shifts in perceived temporal location. *Neurosci Lett* **417**: 219–224.

Jaśkowski, P. (1999). Reaction time and temporal-order judgement as measures of perceptual latency: the problem of dissociations. In G. Aschersleben, T. Bachmann, & J. Müsseler (eds.), *Cognitive Contributions to the Perception of Spatial and Temporal Events* (265–282). North Holland: Elsevier.

Jaśkowski, P., & Verleger, R. (2000). Attentional bias toward low-intensity stimuli: an explanation for the intensity dissociation between reaction time and temporal order judgment? *Conscious Cogn* **9**: 435–456.

Jeffreys, D. A., & Axford, J. G. (1972). Source locations of pattern-specific components of human VEPs. *Exp Brain Res* **16**: 1–21.

King, A. J., & Palmer, A. R. (1985). Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Exp Brain Res* **60**: 492–500.

Kopinska, A., & Harris, L. R. (2004). Simultaneity constancy. *Perception* **33**: 1049–1060.

Lesevre, N. (1982). Chronotopographical analysis of the human evoked potential in relation to the visual field. *Ann NY Acad Sci* **388**: 156–182.

Lewald, J., & Guski, R. (2004). Auditory-visual temporal integration as a function of distance: no compensation for sound-transmission time in human perception. *Neurosci Lett* **357**: 119–122.

Libet, B. (2004). *Mind Time: The Temporal Factor in Consciousness*. Cambridge, MA: Harvard University Press.

Liegeois-Chauvel, C., Musolino, A., & Chauvel, P. (1991). Localization of the primary auditory area in man. *Brain* **114**: 139–153.

Luce, R. D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization.* New York: Oxford University Press.

Macefield, G., Gandevia, S. C., & Burke, D. (1989). Conduction velocities of muscle and cutaneous afferents in the upper and lower limbs of human subjects. *Brain* **112**: 1519–1532.

McKee, S. P., & Smallman, H. S. (1998). Size and speed constancy. In V. Walsh & J. J. Kulikowski (eds.), *Perceptual Constancy* (373–408). Cambridge: Cambridge University Press.

Miyazaki, M., Yamamoto, S., Uchida, S., & Kitazawa, S. (2006). Bayesian calibration of simultaneity in tactile temporal order judgment. *Nat Neurosci* **9**: 875–877.

Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: examining temporal ventriloquism. *Cogn Brain Res* **17**: 154–163.

Navarra, J., Soto-Faraco, S., & Spence, C. (2007). Adaptation to audiotactile asynchrony. *Neurosci Lett* **413**: 72–76.

Navarra, J., Vatakis, A., Zampini, M., Soto-Faraco, S., Humphreys, W., & Spence, C. (2005). Exposure to asynchronous audiovisual speech extends the temporal window for audiovisual integration. *Cogn Brain Res* **25**: 499–507.

Nickalls, R. W. D. (1996). The influence of target angular velocity on visual latency difference determined using the rotating Pulfrich effect. *Vision Res* **36**: 2865–2872.

Pöppel, E. (1988). *Mindworks.* Boston: Harcourt Brace Jovanovich.

Pöppel, E., Schill, K., & von Steinbuchel, N. (1990). Sensory integration within temporally neutral systems states: a hypothesis. *Naturwissenschaften* **77**: 89–91.

Schneider, K. A., & Bavelier, D. (2003). Components of visual prior entry. *Cogn Psychol* **47**: 333–366.

Shore, D. I., Spry, E., & Spence, C. (2002). Confusing the mind by crossing the hands. Brain research. *Cogn Brain Res* **14**: 153–163.

Soto-Faraco, S., Ronald, A., & Spence, C. (2004). Tactile selective attention and body posture: assessing the multisensory contributions of vision and proprioception. *Percept Psychophys* **66**: 1077–1094.

Spence, C., Baddeley, R., Zampini, M., James, R., & Shore, D. I. (2003). Multisensory temporal order judgments: when two locations are better than one. *Percept Psychophys* **65**: 318–328.

Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *J Exp Psychol General* **130**: 799–832.

Spence, C., & Squire, S. (2003). Multisensory integration: maintaining the perception of synchrony. *Curr Biol* **13**: R519–521.

Sternberg, S., & Knoll, R. L. (1973). The perception of temporal order: fundamental issues and a general model. In S. Kornblum (ed.), *Attention and Performance IV* (629–685). New York: Academic Press.

Stone, R. V., Hunkin, N. M., Porrill, J., Wood, R., Keeler, V., & Beanland, M., et al. (2001). When is now? Perception and simultaneity. *Proc Roy Soc Lond B* **268**: 31–38.

Sugita, Y., & Suzuki, Y. (2003). Audiovisual perception: implicit estimation of sound-arrival time. *Nature* **421**: 911.

Vatakis, A., & Spence, C. (2006). Audiovisual synchrony perception for speech and music assessed using a temporal order judgment task. *Neurosci Lett* **393**: 40–44.

von Békésy, G. (1963). Interaction of paired sensory stimuli and conduction in peripheral nerves. *J Applied Physiol* **18**: 1276–1284.

von Békésy, G. (1967). *Sensory Inhibition.* Princeton, NJ: Princeton University Press.

Vroomen, J., Keetels, M., de Gelder, B., & Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Cogn Brain Res* **22**: 32–35.

Walsh, V., & Kulikowski, J. (1998). *Perceptual Constancy: Why Things Look as They Do*. Cambridge: Cambridge University Press.

Wilson, J. A., & Anstis, S. M. (1969). Visual delay as a function of luminance. *Am J Psychol* **82**: 350–358.

Zampini, M., Brown, T., Shore, D. I., Maravita, A., Roder, B., & Spence, C. (2005). Audiotactile temporal order judgments. *Acta Psychol (Amst)* **118**: 277–291.

Zampini, M., Guest, S., Shore, D. I., & Spence, C. (2005). Audio-visual simultaneity judgments. *Percept Psychophys* **67**: 531–544.

Zampini, M., Shore, D. I., & Spence, C. (2005). Audiovisual prior entry. *Neurosci Lett* **381**: 217–222.

16

Relative timing and perceptual asynchrony

DEREK H. ARNOLD

Summary

How do human observers determine the relative timings of different events? One perspective, which I shall refer to as the brain–time theory of perception, suggests that apparent timing is related to when specific analyses are *concluded* within distinct and relatively independent regions of the brain. This proposal is controversial, not least because it suggests that temporal perception is error prone and subject to the rates at which analyses are concluded in different parts of the brain. One observation that may favor this perspective is that physically coincident changes in color and direction can appear asynchronous – a perceptual asynchrony. In this chapter I will review the theoretical interpretations and empirical evidence that relate to this phenomenon. I will argue that this timing illusion provides good evidence for a relationship between the time courses of sensory processing in the brain and perceived timing.

16.1 Introduction

How do we determine relative timing? Human observers can determine the relative timings of a remarkable variety of events. For instance, we can judge the timings of visual relative to other visual (Moutoussis & Zeki 1997a,b), auditory (Fujisaki et al. 2004), and haptic (Vogels 2004) events. Subjectively it seems that one of the *events* can be entirely intrinsic to the nervous system, like the sensation of reaching seven while mentally counting from one to ten. The fact that these judgments can be made dictates that the necessary information is encoded in a form that can then be reported – but it is not clear how this feat is achieved.

Independent processes and brain structures encode different perceptual events, so determining relative timing may be inherently difficult. Although there is some debate concerning the degree to which these statements are true of vision (Lennie 1998), a consideration of clinical (Zihl et al. 1983; Cowey & Heywood 1997) and neurophysiological (Zeki 1978; Livingstone & Hubel 1988) evidence suggests that judgments of visual coincidence may often necessitate the comparison of events encoded in different regions – and possibly at different times – in the brain.

Before we go further, some terminology should be specified to avoid confusion. *Relative perceptual timing* refers to the apparent timing of one perceptual event relative to another. This issue differs from that of *apparent time* – being the time at which an event seems

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

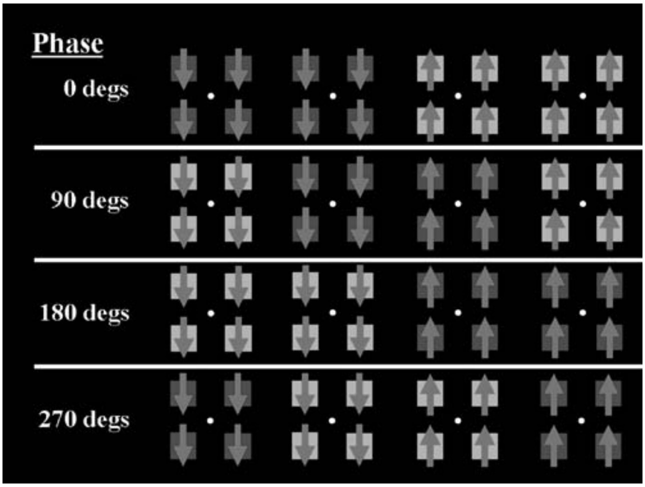


Fig. 16.1 A depiction of four phasic relationships between color and motion. At a phasic relationship of 0 deg, the color green is always paired with upward motion and red with downward. At a phasic relationship of 180 deg, this situation is reversed. At phasic relationships of 90 deg and 270 deg upward and downward motions are paired equally with red and green – no physical correlation exists between color and motion at these phasic relationships.

to have occurred. All data discussed here will relate to relative perceptual timing and the insights this provides into the process(es) by which we determine perceptual coincidence. For now I shall leave it to others to determine *when* an event seems to occur and whether or not this moment lags the physical timing of that event (see Libet et al. 1983; Libet 1985; van de Grind 2002).

Recently a number of studies have addressed the issue of relative timing in relation to a striking perceptual phenomenon – an apparent perceptual asynchrony between color and motion (Moutoussis & Zeki 1997a,b; Arnold et al. 2001; Johnston & Nishida 2001; Viviani & Aymoz 2001; Arnold & Clifford 2002; Nishida & Johnston 2002; Bedell et al. 2003; Paul & Schyns 2003; Adams & Mamassian 2004; Clifford et al. 2004; Enns & Oriet 2004; Moradi & Shimojo 2004; Arnold 2005). The stimulus used in the initial series of experiments (Moutoussis & Zeki 1997a,b) alternated in color (between being red and green) and direction of motion (between up and down). The relative timings of the changes in color and direction were manipulated (see Fig. 16.1), and observers were required to indicate which colors seemed to coexist with which directions of motion. Surprisingly, observers were most likely to report that a color was moving in a given direction when the change in color lagged the change in direction by ∼80 msec.

These findings seem to suggest that perceptual events that occur at the same physical time can seem to occur at different times. This was interpreted as being indicative of a temporal hierarchy of sensory processing, with changes in color being processed more rapidly than changes in direction (Moutoussis & Zeki 1997a,b). The central assumption underlying this interpretation is that the time course of sensory analyses will influence relative timing judgments – so I will refer to this proposal as the brain time hypothesis.

The initial interpretation of the apparent perceptual asynchrony between color and motion (Moutoussis & Zeki 1997a,b) may not have attracted the attention that it perhaps deserves because the findings are seemingly at odds with physiological evidence showing that direction changes can begin to be processed in cortex before color changes (Schiller & Malpeli 1978; Munk et al. 1995). However, this contradiction may be more apparent than factual. At this point we do not know how the brain encodes relative timing and thus we cannot be sure how timing perception maps onto cortical activity. It is at least possible that the crucial factor is not when neural activity can first be detected at the level of a single neuron but by when population level analyses are concluded. These might typically be completed more rapidly for color than they are for motion. Obviously this is entirely speculative and it would be much preferable to know *how* timing perception relates to brain activity. However, the brain time hypothesis does provide a simple and falsifiable prediction – if the time course of sensory processing is changed, so too will our sense of timing.

The initial demonstration of color/motion perceptual asynchrony has inspired a number of subsequent studies. Some have argued for slight modifications of the *brain time* interpretation (Arnold et al. 2001; Viviani & Aymoz 2001; Arnold & Clifford 2002; Bedell et al. 2003; Paul & Schyns 2003; Adams & Mamassian 2004; Clifford et al. 2004; Arnold 2005). Others are more critical and have argued in favor of alternate proposals. Instead of a perceptual asynchrony driven by different processing times (Moutoussis & Zeki 1997a,b) it has been argued that the perceptual asynchrony occurs because of a postdicitve analysis that determines the perceptual properties of new surfaces (Moradi & Shimojo 2004), an error-prone process of temporal marker matching (Johnston & Nishida 2001; Nishida & Johnston 2002), or attention switching (Enns & Oriet 2004). I will discuss each of these developments in turn.

16.2 Modified brain time

16.2.1 Time scale

In the initial perceptual asynchrony studies (Moutoussis & Zeki 1997a,b) observers were required to indicate predominant pairings between either color and motion direction, between color and orientation, or between orientation and motion direction. In these sorts of tasks it is likely that there will be some situations in which the observer is aware that a given attribute is paired with both states of the other attribute. For instance, at phasic relationships of 90 deg and 270 deg (see Fig. 16.1) each direction of motion (up and down) is paired for equal periods of time with each color (red and green). The observer may be aware of this situation but would still be required to pair a given direction with just one of the two colors (Moutoussis & Zeki 1997a,b). It is not clear, therefore, if the temporal bias revealed by this task is related to coincidence detection or to a temporal bias that emerges when trying to determine predominant pairings over extended periods of time.

Subsequent studies have used similar stimuli in that they contained different stimulus attributes alternating between different states (as depicted in Fig. 16.1), but have used different perceptual tasks. One of these used a stimulus containing alternations of color and orientation (Clifford et al. 2003). Observers were required to complete two tasks: (1) To indicate when the stimulus changes appeared to be simultaneous – a different task to that used in the initial studies (Moutoussis & Zeki 1997a,b). (2) To indicate predominant perceptual pairings, even if none seemed to exist – the same task as the initial studies (Moutoussis & Zeki 1997a,b). Only the latter task revealed a temporal bias, the former showed that physically synchronous color and orientation changes also seemed to be perceptually synchronous (Clifford et al. 2003).

These data suggest an important distinction between temporal judgments at different time scales. Synchrony judgments concern the coincidence of two events at a micro time scale whereas pairing judgments concern predominant pairings over extended durations. Judgments over these time scales can reveal different temporal biases, an observation that implicates dissimilar perceptual processes (Bedell et al. 2003; Clifford et al. 2003). Here I am primarily interested in the micro time scale and the insights this provides into mechanisms of coincidence detection. As the color/orientation perceptual asynchrony only occurs over extended time scales (Moutoussis & Zeki 1997b; Clifford et al. 2003), it is unlikely to be causally related to different processing times or to clarify the mechanisms underlying coincidence detection at micro time scales (Clifford et al. 2003).

Another study has used synchrony judgments to examine alternations in color and motion direction (Nishida & Johnston 2002). It was found that, in some circumstances (see temporal marker discussion below), physically delayed color changes did appear to coincide with earlier direction changes – consistent with the findings of Moutoussis and Zeki (1997a,b). This observation, of an illusory temporal conjunction between color and direction changes, suggests that the method by which we determine perceptual coincidence is error prone. Identifying the cause/s of this error may provide considerable insights into the mechanisms of perceptual timing.

16.2.2 No fixed temporal hierarchy

Although the initial study was restricted in scope to the contexts of color and motion (Moutoussis & Zeki 1997a), a follow-up study also included orientation (Moutoussis & Zeki 1997b). Using the same methodology it was found that observers tended to pair specific colors with *earlier* instances of orientation (~63 msec), and with even *earlier* instances of motion (~118 msec). These data seemed to be interpreted in terms of a fixed temporal hierarchy of distinctive processing times for different stimulus properties, with changes in color being processed faster than changes in orientation, which were processed faster than changes in motion direction (Moutoussis & Zeki 1997b).

The link between perceptual asynchrony and fixed distinctive processing times has been undermined by studies showing that the magnitude of perceptual asynchrony is variable. For instance, both Arnold and Clifford (2002) and Bedell et al. (2003) have shown that the magnitude of perceptual asynchrony between color and motion varies as a function of the angular difference between the contrasted directions. The asynchrony is greatest if opposite directions are contrasted and decreases as the angular difference between the alternating directions is reduced (Arnold & Clifford 2002; Bedell et al. 2003). The magnitude of color/motion perceptual asynchrony is also influenced by the salience of the contrasted stimulus attributes (Adams & Mamassian 2004) and by the luminance (Bedell et al. 2006). All these manipulations can reasonably be expected to influence the time course of sensory processing. Direction change magnitude can influence the time course of sensory processing because of inhibition (Barlow & Levick 1965; Snowden et al. 1991; Bair et al. 2002). Similarly, salience and luminance can influence the time course of sensory processing because increasing stimulus strength can shorten neural latencies (Munk et al. 1995; Gawne et al. 1996).

The variability of color/motion perceptual asynchrony cautions against attributing the phenomenon to fixed distinctive processing times for different stimulus properties (Arnold & Clifford 2002; Bedell et al. 2003; Adams & Mamassian 2004; Bedell et al. 2006) – but this variability is not inconsistent with the brain time account. The major implication of brain time is that the time courses of sensory processing and perceptual experience are related. This does not imply that the processing time for any type of perceptual analysis will be constant. Rather, analyses are likely to require variable periods of time to complete – reflecting the computational demands of specific analyses and the adaptive state of the processing system (see Bedell et al. 2003 and Ogmen et al. 2004 for similar arguments).

16.2.3 Perceptual asynchrony depends on perceiving direction changes

In some circumstances, motions that are separated in both time and space can become grouped to create an impression of a persistent moving surface. Consider the situation in which direction changes are signaled by multiple dots. When all the dots change between the same directions at the same times, observers experience a single field of coherently moving dots. Another situation prevails when the relative timings of the changes are randomized – such that the individual dots still alternate between different motion directions but the timings of the changes between those directions are randomized. In this situation observers tend to see two unchanging transparent surfaces that contain dots drifting in different directions (Clifford et al. 2004; Kanai et al. 2004).

An interesting dissociation emerged when these two types of moving stimuli were contrasted. When all dots changed directions at the same times, such that a single coherently moving surface was experienced, observers tended to pair direction changes with delayed color changes – the typical color/motion perceptual asynchrony (Clifford et al. 2004). However, when the timings of direction changes were random, such that observers experienced two persistent transparent surfaces, pairings of color and motion were veridical (Clifford et al. 2004).

Findings relating motion transparency to perceptual pairings of color and motion therefore pose an important caveat for the brain time account of perceptual asynchrony: observers will only misbind motion with later instances of color when a persistent surface appears to change direction (Clifford et al. 2004; Moradi & Shimojo 2004). Here it is interesting to note that processing in the human visual brain region V5 has been strongly implicated as a determinant of motion perception (Zeki et al. 1993; Tootell et al. 1995). It has also been suggested that motion transparency occurs when the density of moving elements is sufficient to ensure that there is always movement in specific directions within the relatively large receptive fields of V5 neurons. Hypothetically this would ensure that motion signals from V5 are constant and that perceived direction does not change (Qian & Andersen 1994; Qian et al. 1994a,b; Clifford et al. 2004). The influence of motion transparency on color/motion perceptual asynchrony may therefore indicate that the dynamics of activity in V5 is a strong determinant of the color/motion perceptual asynchrony (Clifford et al. 2004).

16.2.4 Section summary

Illusory temporal conjunctions may provide an invaluable tool with which to explore the mechanisms and processes underlying perceptual timing. Studies using simultaneity judgments have found no evidence for perceptual asynchrony between color and orientation – so the comparison of these attributes is unlikely to be greatly informative in this context. However, temporally offset color and direction changes are often judged as being perceptually synchronous. Interestingly, color/motion perceptual asynchrony only occurs when a persistent surface appears to change direction – suggesting that the processing delays involved in bringing about a change in perceived direction may play a causal role in the generation of this illusion. The magnitude of color/motion perceptual asynchrony is variable, which cautions against attributing the phenomenon to fixed and distinctive processing time differences. However, the magnitude of color/motion perceptual asynchrony varies in a manner that is broadly consistent with the known dynamics of sensory processing. Thus, although it cannot be attributed to fixed differences, it is entirely plausible that color/motion perceptual asynchrony arises because of a close relationship between the dynamic time course of sensory processing and perceived timing.

16.3 Temporal marker matching

Instead of processing time differences, Nishida and Johnston (2002) have suggested that the color/motion perceptual asynchrony happens because of an error-prone process of temporal marker matching. According to this proposal, markers are attached to salient temporal events. The terminology used to describe temporal events refers to transitions (which can be detected by sampling two points in time) as being first-order events and to turning points (which can only be detected by sampling three points in time) as being second-order events. In this vernacular, color and position changes are first-order events, whereas direction reversals are second-order events (Nishida & Johnston 2002). According

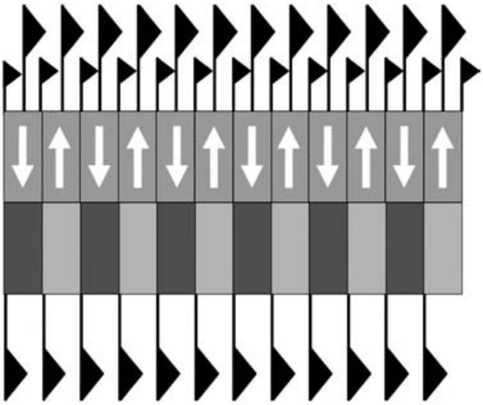


Fig. 16.2 A graphic description of the temporal marker matching account of perceptual asynchrony. According to the proposal, temporal markers are attached to different types of temporal event. Two hypothetical event types are proposed: first-order changes, which can be detected by sampling two points in time, and second-order changes, which can only be detected by sampling three points in time. Direction reversals would therefore be second-order changes and color and position changes would be first order. Here the large flags signify markers attached to first-order events, and the smaller flags signify markers attached to second-order events. Perceptual asynchrony is supposed to occur because markers attached to first-order position changes are inappropriately matched with markers attached to temporally offset color changes.

to the proposal, second-order events are difficult to detect when they are embedded within rapid alternations (∼1 Hz and faster, see Fig. 16.2). Theoretically the temporal markers attached to first-order position changes (which separate the direction reversals) are more salient than the temporal markers attached to second-order direction reversals and, as a consequence, the temporal markers attached to first-order color changes are erroneously paired with the temporal markers attached to first-order position changes.

It is far from clear how the temporal marker account, at least in its original form (Nishida & Johnston 2002), could account for pre-existing findings related to color/motion perceptual asynchrony. For instance, if perceptual asynchrony is driven by mismatching markers attached to temporally offset events (Nishida & Johnston 2002) asynchrony magnitude should vary with the size of the average temporal offset between the mismatched events. This would be determined by the duration of the alternation cycle – longer cycles should cause larger apparent asynchronies. However, it has been shown that alternation rates can be changed without influencing perceptual asynchrony magnitude (Moutoussis & Zeki 1997a,b; Bedell et al. 2003).

Perhaps a more dramatic challenge for the temporal marker account is that it does not appear to predict the direction of perceptual asynchrony. Color/motion perceptual asynchrony is characterized by a tendency to pair direction changes with relatively delayed color changes. A relative processing delay, with color changes being processed more rapidly than direction changes (Moutoussis & Zeki 1997a,b), could readily explain both the consistency and direction of this effect. However, the temporal marker theory, at least in its original form, does not explain why there is a tendency to pair markers attached to first-order position transitions with markers attached to delayed rather than to earlier color changes (Nishida & Johnston 2002).

The emphasis of temporal marker theory on mismatching first- and second-order temporal events also seems to be problematic. Both Arnold and Clifford (2002) and Bedell et al. (2003) have shown that the magnitude of color/motion perceptual asynchrony varies according to the angular difference between the contrasted directions of motion. Using the vernacular of temporal marker theory, this manipulation does not change the types of temporal events being compared – direction changes, be they big or small, can only be detected by sampling three points in time and should therefore be characterized as second-order temporal events (Nishida & Johnston 2002). The fact that asynchrony magnitude varies when temporal event types are held constant implies that the color/motion perceptual asynchrony cannot comprehensibly be explained by temporal marker theory (Nishida & Johnston 2002).

16.3.1 Modified temporal marker matching

Recently the temporal marker theory has evolved to acknowledge the influence of the angular difference between contrasted directions of motion. First, the authors replicated the basic finding concerning asynchrony magnitude and motion direction angular differences (Amano et al. 2007). Previously such findings have been taken as strong evidence for the brain time account. Amano et al. (2007), however, interpreted these data within a temporal marker framework. Specifically, they proposed that the angular difference between the contrasted directions of motion had influenced the time course of sensory processing and that this, in turn, had influenced the placement of temporal markers. This interpretation seems problematic. Most importantly, it effectively transforms the temporal marker account of perceptual asynchrony into another form of brain time.

Perhaps the most attractive conceptual feature of the original temporal marker account was that it strove to dissociate perceived timing from the dynamic, and therefore unreliable, time course of sensory processing (Johnston & Nishida 2001; Nishida & Johnston 2002). Recent developments have greatly clarified the temporal marker account but seem to have undermined this conceptual advantage. According to the modified version of the temporal marker account (Amano et al. 2007), perceived timing is determined by attaching temporal markers to initial, as opposed to completed, sensory analyses in cortex. According to this proposal, sensory processing *within* cortex should have little impact on perceived timing – but our sense of timing would still be subject to brain time. Primarily this would be true because the timings of initial cortical responses are subject to variable delays due to sensory processing in subcortical structures, such as the retina and the LGN.

The revised version of the temporal marker account seems to be a development of an earlier proposal, that perceived timing is related to the timing of the first cortical response to a sensory event (Libet et al. 1979, 1985; Libet et al. 1983; see van de Grind 2002, for a critical commentary). According to this perspective, timing perception is subject to brain time in that illusions can be instigated by inducing differential delays between physical exposure to a stimulus and the onset of activity in cortex (Roufs 1963; Wilson & Anstis 1969). However, by attributing perceived timing to when cortical analyses *commence*, rather than when they *finish*, the influence of processing time differences *within* cortex should be minimized (Libet et al. 1979, 1985; Libet et al. 1983; Amano et al. 2007). Thus timing illusions might occur, but they should not become further exacerbated by sensory processing within cortex.

It seems that the revised version of the temporal marker account could be conclusively falsified by demonstrating that processing time differences arising because of processing *within* cortex can induce a subjective perceptual asynchrony. This would have to be done in a context that precluded processing prior to the cortex from contributing to the asynchrony. A note of caution should be sounded here though. It is unclear what, precisely, is meant by an initial cortical response. If this were considered to be the point in time at which a population response in cortex can be distinguished from noise, considerable delays could still occur because of processing *within* cortex. However, this proposal would seem to go against the conceptual spirit of the temporal marker account as such a process would be dependent upon the completion of an analysis within cortex (the process by which a sensory signal is distinguished from neural noise). It would also be particularly difficult to distinguish this latter proposal from other explicitly brain time accounts (Moutoussis & Zeki 1997a,b; Arnold & Clifford 2002; Bedell et al. 2003).

16.3.2 Section summary

It has been suggested that the color/motion perceptual asynchrony might be driven by mismatching different types of events. Events detectable by sampling two points in time are termed first order and those that can only be detected by sampling three points in time are second order. As yet it is not clear how this characterization could explain the direction or relatively consistent magnitude of the color/motion perceptual asynchrony. More recently the temporal marker account has been revised to acknowledge the influence of the variable processing times that can be instigated by contrasting directions separated by differing angles. This development transforms the temporal marker account into another form of the brain time account. However, it still does not address *why* color/motion perceptual asynchrony has a consistent direction and magnitude. Thus, at this point, temporal marker theory does not appear to provide a clear and/or comprehensive account of color/motion perceptual asynchrony.

16.4 Motivations underlying temporal marker matching

Although temporal marker theory does not appear to provide a clear or comprehensive account of color/motion perceptual asynchrony, the motivations that drove the formation of the theory should be addressed. These included (1) that color/motion perceptual asynchrony does not cause reaction time differences, which one might expect if analyses of color are completed more rapidly than analyses of direction change; (2) that color/motion perceptual asynchrony only occurs for rapid stimulus alternations, suggesting that the perceptual asynchrony might only occur in special circumstances; (3) that color/motion perceptual asynchrony is predicted by temporal structure rather than attribute type, suggesting that an emphasis on the former might explain the illusion. I shall address each of these observations in turn.

16.4.1 Color/motion perceptual asynchrony is not matched to reaction time differences

Reaction times may seem like a simple measure of perceptual processing, but they are quite complicated. In addition to any effects of sensory processing, they are influenced by both the preparation and execution of motor responses. Thus even when a reaction time difference is observed, it is unclear if that difference can be attributed to sensory or to motor-related processing. Further complicating matters is the possibility of compensation. For instance, delayed sensory processing could be compensated for by more rapid motor processing, or vice versa. The implications of a null result in this context are therefore also unclear.

Even if a reaction time measure could reliably be attributed to sensory processing, it would be unclear if the sensory processing being tapped was necessarily related to the determination of perceived timing. As sensory processing is multifaceted and distributed across different brain structures, it is at least possible that a processing difference could affect reaction times without impacting the determination of perceived timing.

In summary, the dynamics of perceptual experience, our sensations of relative timing, and the speed by which we can react to a physical event might all be unrelated. As a consequence, reaction time data do not pose a great conceptual dilemma for any account of perceived timing.

16.4.2 Color/motion perceptual asynchrony only occurs for rapid stimulus alternations

Nishida and Johnston (2002) found that the color/motion perceptual asynchrony only occurs for rapid stimulus alternations, $\sim$1 Hz and faster. Although this may suggest an absence of processing differences at slow alternation rates, it does not prove that perceptual asynchrony observed at high alternation rates is not related to processing differences. Nishida and Johnston (2002) recognized this and sought to support the finding with the additional observations that temporal order judgments for specific color and direction changes were accurate, even when said changes were embedded in rapid stimulus alternations. However, there have been a number of failures to replicate these results. Viviani and Aymoz (2001), and then Aymoz and Viviani (2004) found that temporal order judgments were inaccurate when comparing single color changes and motion onsets – color changes seemed to precede synchronous motion onsets. Adams and Mamassian (2004) found that single color changes and direction reversals were similarly offset. However, Bedell et al. (2003) found that temporal order judgments were accurate when comparing single color changes and direction reversals – even when the stimulus alternation was rapid.

The variability of color/motion temporal order judgments, and the dependence of perceptual asynchrony on alternation rate, suggests the possibility of multiple temporal cues – at least one of which prompts accurate timing and another that does not. Nishida and Johnston (2002) suggested that the variability arises because, at rapid alternations, second-order temporal markers are more difficult to detect. But given the ill-defined nature of temporal markers and the fact that asynchrony magnitude varies when the type of hypothetical marker is held constant (Arnold & Clifford 2002; Bedell et al. 2003), it is just as probable that the variability reflects changed processing dynamics or the influence of different perceptual processes.

One suggestion that could explain the variability of color/motion timing judgments is that, in some circumstances, relative timing could be signaled by neural offset transients. Offset transients are shorter and less variable than are onset transients and could therefore provide a more reliable signal for perceptual timing (Bair et al. 2002). However, neural offsets could only signal the starting point of a stimulus transition – they could play no role in signaling visual content. Perhaps the variability of color/motion timing judgments reflects two perceptual strategies – one that monitors the stimulus for offset transients that signal the starting points for stimulus transitions and one that monitors the stimulus for content changes, which are therefore susceptible to processing delays.

Color/motion timing judgments are variable (Viviani & Aymoz 2001; Nishida & Johnston 2002; Bedell et al. 2003; Adams & Mamassian 2004; Aymoz & Viviani 2004). This variability cannot be predicted on the basis of the type of temporal event. Sometimes the relative timings of color changes (hypothetically first-order events) and direction reversals (second-order events) are judged accurately (Nishida & Johnston 2002; Bedell et al. 2003) and sometimes not (Viviani & Aymoz 2001; Adams & Mamassian 2004; Aymoz & Viviani 2004). Although it may take some time to clarify this tangle of apparently contradictory findings, it seems reasonable to suggest that apparent discrepancies may be driven by the provision of different temporal cues. It should be stressed that the variability of color/motion perceptual coincidence does not refute the possibility that perceptual asynchrony can be induced by processing differences (Moutoussis & Zeki 1997a,b).

16.4.3 Color/motion perceptual asynchrony is predicted by temporal structure rather than attribute type

Inspired by their emphasis on temporal event type, Nishida and Johnston (2002) attempted to demonstrate that the direction and magnitude of the color/motion perceptual asynchrony could be predicted by this characterization. To this end they contrasted combinations of first- and second-order changes in color and position. First-order position changes resembled an apparent motion stimulus, generated by changing the phase of the component waveforms in a plaid stimulus, +/– 0.25 cycle, every 250 msec. A moving plaid, changing direction every

250 msec, was used to signal second-order position changes. First-order color changes were signaled by abrupt red/gray color changes, whereas second-order color changes were signaled by gradual transitions from gray to red, then from red to gray.

When gradual color changes (second order) were coupled with abrupt position changes (first order), observers tended to pair colors with relatively delayed positions (Nishida & Johnston 2002). If abrupt position changes are directly analogous to motion, this would constitute a reversal of the normal color/motion perceptual asynchrony. It does not seem unreasonable, however, to suggest that the reversal of the typical color/motion asynchrony in this case is driven by processing differences. The very different types of position change may tap different visual processes with different temporal dynamics. Alternatively, they may drive a common process in very different ways – direction reversals may prompt greater direction inhibition than do abrupt position changes. Similar statements can be made in relation to the very different types of color change. The observation that different stimuli can prompt different temporal relationships is entirely consistent with the suggestion that visual processing can influence the time course of sensory experience and therefore apparent timing (Moutoussis & Zeki 1997a,b).

Although Nishida and Johnston (2002) focused on the combination of first-order position and second-order color changes, of equal interest are the combinations of first-order color and position and of second-order color and position changes. When these combinations were compared, and the theoretical types of temporal event were constant, positions tended to be paired with slightly delayed colors (perceptual asynchronies of $\sim$25–55 msec). These findings suggest that an emphasis on temporal event type cannot provide a comprehensive account of color/motion perceptual asynchrony.

16.4.4 Section summary

The findings that inspired the temporal marker account of perceptual asynchrony do not refute the brain time account. Reaction time measures may not tap the dynamics of perception or perceived timing, and changing the physical properties of the stimulus may change the dynamics of perceptual processing. Moreover, the color/motion perceptual asynchrony cannot be predicted by mismatching theoretically different event types because asynchrony magnitude can change when event types are held constant.

16.5 Postdictive analysis

The brain time account of color/motion perceptual asynchrony argues that there is a relationship between the time courses of sensory processing and perceptual experience (Bartels & Zeki 1998; Zeki 2002). However, it has been argued that neural processing times may not correlate directly with relative perceptual timing (Dennett & Kinsbourne 1992; Johnston & Nishida 2001; Rao et al. 2001; Nishida & Johnston 2002; Moradi & Shimojo 2004). According to this view there is a need to resolve ambiguities of timing that might arise because of the variable time courses of neural processing. Instead of relative timing being determined by when sensory analyses are completed, it has been argued that timing might be determined by an interpretive analysis that compensates for the variability of neural transmissions (Dennett & Kinsbourne 1992; Johnston & Nishida 2001; Rao et al. 2001; Nishida & Johnston 2002; Moradi & Shimojo 2004).

One way of compensating for the variability of neural processing would be to attribute the timing of stimulus changes to the points in time at which analyses *commence* rather than to the points in time at which they *finish*. This form of analysis has been described as postdictive, as perceptual experience is supposed to be delayed, thereby allowing the visual system to take into account information from after an event before committing to a visual interpretation (Eagleman & Sejonowski 2000; Rao et al. 2001; Moradi & Shimojo 2004). It has been argued that the color/motion perceptual asynchrony might be caused by a postdictive analysis (Moradi & Shimojo 2004).

This interpretation posits that when a new surface appears, or when a persistent surface changes direction, an analysis of the surface's properties is triggered – including an analysis of its color (Moradi & Shimojo 2004). Hypothetically, this analysis persists for a period of time but the analysis outcomes are experienced as having originated from the beginning of the analysis. How could this process cause color/motion perceptual asynchrony? Well, during the analysis period color could change – say from red to green. The analysis would then include a brief instance of red and a more prolonged instance of green. Due to averaging, the perceived color of the new surface would be green – thereby inappropriately pairing a temporally offset color change with an earlier direction reversal (Moradi & Shimojo 2004). The potential for illusory temporal conjunctions being created by this hypothetical process is exacerbated by the assumption that color is not treated consistently during analysis periods. Instead, it is assumed that perceptual outcomes are biased in favor of the later stages of analysis (Moradi & Shimojo 2004; see Fig. 16.3 for a graphic description).

The postdictive account of color/motion perceptual asynchrony (Moradi & Shimojo 2004) is similar to the modified version of the temporal marker matching account (Amano et al. 2007). However, there are some important differences. According to the modified temporal marker matching account (Amano et al. 2007) all sensory events are temporally marked, so they should seem to have occurred at the time of the first cortical response elicited by the event (Libet et al. 1979, 1985; Libet et al. 1983; Amano et al. 2007). Thus, unless a subcortical processing difference is assumed, the account does not appear to be able to predict the direction or relatively consistent magnitude of the color/motion perceptual asynchrony. In contrast, the dictive account of color/motion perceptual asynchrony (Moradi & Shimojo 2004) suggests that postdictive analyses are triggered by direction reversals. This would explain the consistent bias to pair color with earlier instances of motion.

The brain time and postdictive accounts of perceptual asynchrony predict different optimal conditions for making temporal judgments. According to the brain time account (Moutoussis & Zeki 1997a,b), direction changes are paired with later color changes because it takes longer to process direction changes – so the optimal conditions for pairing color and direction changes should occur when direction changes are delayed relative to color changes. According to the postdictive account (Moradi & Shimojo 2004) direction changes

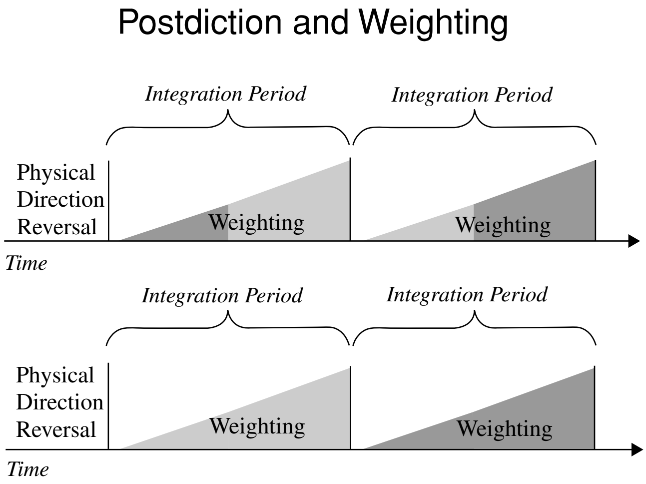


Fig. 16.3 Schematics depicting the theoretical postdiction and temporal averaging account of color/motion perceptual asynchrony. It is assumed that physical direction reversals (vertical lines) trigger a process of temporal averaging by which the apparent color of the moving surface is determined. The perceptual outcome of the analysis is postdictive as it is assumed that the average color is experienced as having persisted from when the analysis commenced. It is also assumed that color is not treated equally during the analysis – outcomes are biased in favor of colors present during the later stages. Because of this, the apparent colors of the moving surfaces depicted above and below will both be green and then red – even though the surface depicted above is both red and green for equal periods during the analysis periods triggered by the direction reversals.

trigger an extended process of temporal averaging during which the apparent color of the moving surface is determined. The color/motion perceptual asynchrony is supposed to occur because the perceptual outcome of this extended period of analysis (color) is experienced as having persisted from the point in time at which the direction change occurred. Therefore, according to this proposal, the optimal conditions for pairing color and direction changes should occur when direction and color changes are physically coincident. This will maximize the proportion of the analysis period containing a single color – so the outcome of the averaging process should not be at all ambiguous.

Moradi and Shimojo (2004) recognized that one of the central predictions of the brain time account is that color/motion pairing should be facilitated by delaying color relative to direction changes. They therefore examined the effects of delaying color changes and found that this did not facilitate color/motion pairing (Moradi & Shimojo 2004). However, they used fairly rapid alternation rates (between 3.57 and 5.3 Hz) at which the perceptual pairing of color and motion was impossible for their observers (see Moradi & Shimojo 2004, Fig. 16.3). Thus it did not matter if color changes were delayed or not – the task was still impossible. If we want to see if inducing color change delays can facilitate color/motion pairing we need to examine conditions in which it is at least sometimes possible to complete the task.

I have recently examined this issue in two ways (Arnold 2005). At a fixed rate of alternation (1.25 Hz) I found that color/motion pairing is far more systematic when color changes are delayed by 80 msec as opposed to when they are physically synchronous. When color and direction changes were physically synchronous, pairing performance did not differ from chance (Arnold 2005). This observation is also apparent upon close inspection of data from previous studies (Moutoussis & Zeki 1997b; Nishida & Johnston 2002). I also demonstrated that observers are able to match colors and motion directions at faster alternation rates when color changes are delayed by 120 msec relative to direction changes. When color changes lagged direction changes, observers could pair colors and motion directions at alternation rates of ∼2.82 Hz, whereas they could only pair these attributes at rates of ∼1.97 Hz when the changes were physically synchronous (Arnold 2005).

Color/motion perceptual pairing can therefore be facilitated by inducing a time lag between the changes in direction and the changes in color (Arnold 2005). Although this observation is predicted by the brain time account of perceptual asynchrony (Moutoussis & Zeki 1997a,b), it seems inconsistent with the postdictive account (Moradi & Shimojo 2004). This apparent inconsistency might be resolved if there was a delay between detection of the triggering event and the start of the period of postdictive analysis. For instance, a postdictive analysis might be triggered by the detection of a direction reversal, but the postdictive analysis of color may not begin until sometime later.

This suggestion differs from earlier postdictive hypotheses (Eagleman & Sejnowski 2000; Rao et al. 2001). The major computational advantage of the initial suggestions was that time-consuming analyses would have little or no impact upon timing judgments. In contrast, the proposal above suggests that timing errors would occur as a direct consequence of a time-consuming neural process – the delay between a direction reversal being detected and the beginning of a postdictive analysis. This proposal therefore lacks the functional advantages of earlier postdictive hypotheses (Eagleman & Sejnowski 2000; Rao et al. 2001), thus there seems to be no reason to prefer this hypothesis in comparison to the brain time account of perceptual asynchrony.

16.5.1 Section summary

It has been suggested that the color/motion perceptual asynchrony might be driven by a postdictive analysis. According to this proposal direction changes trigger a prolonged averaging process during which the apparent color of the moving surface is determined. Two factors are supposed to cause perceptual asynchrony: (1) Analysis results are experienced as having persisted from the beginning of the analysis. (2) The apparent color determined by averaging can differ from the color that existed at the beginning of the analysis. Accordingly, the best conditions for pairing color and motion should occur when color and direction changes are physically synchronous and the analysis period contains just one color. This is not true. Observers are better at pairing color and motion when changes in color lag changes in direction – an observation that is entirely consistent with the brain time account of perceptual asynchrony but is inconsistent with the postdictive account.

16.6 Attention switching

It has been suggested that perceptual asynchrony might arise because of a delay between redirecting attention from a “defining” attribute to a “report” attribute. The idea is that the observer needs to first detect a change in one feature attribute (say a direction change) and then redirect attention to judge the state of another feature (color). Support for this perspective is garnered from the observation that switching feature roles (defining/report) can reverse the apparent direction of perceptual asynchrony. For instance, if observers are instructed to report the color of a vertical bar they tend to pair orientation with delayed colors – as demonstrated by Moutoussis and Zeki (1997b). However, if observers are instructed to report the orientation of red they tend to pair color with delayed orientations (Enns & Oriet 2004).

Another study used the same manipulation to examine the influence of attention on perceptual asynchrony but failed to find a reversal of the normal effect (Clifford et al. 2003). However, this study did find that perceptual asynchrony magnitude could be modulated by attention (Clifford et al. 2003). The two studies therefore disagree in terms of effect magnitude. In one, attention effects were sufficient to reverse the direction of apparent perceptual asynchrony (Enns & Oriet 2004), whereas in the other they were not (Clifford et al. 2003). However, the two studies are in agreement concerning the fact that attention can influence temporal judgments.

The effects of attention on temporal judgments have long been acknowledged. Typically, attended changes appear to precede unattended changes (Sternberg & Knoll 1973; Reeves & Sperling 1986). Recognizing this, experimenters have often taken precautionary measures when assessing perceptual asynchrony. Some experimenters have avoided using “defining” and “report” attributes by requiring observers to make forced choice judgments concerning the simultaneity (Nishida & Johnston 2002; Bedell et al. 2003; Clifford et al. 2003; Arnold 2005) or temporal order (Viviani & Aymoz 2001; Nishida & Johnston 2002; Adams & Mamassian 2004; Aymoz & Viviani 2004) of stimulus changes. In all of these studies apparent perceptual asynchronies were observed even though observers were not instructed to attend to any specific attribute.

When talking about attention, it is often beneficial to draw a distinction between endogenous and exogenous effects. Endogenous effects refer to the observer’s ability to selectively attend to a specific stimulus. Some of the studies described previously have shown that endogenous attention can at least modulate perceptual asynchrony magnitude (Clifford et al. 2003; Enns & Oriet 2004). However, the studies that have avoided using “defining” and “report” features have shown that perceptual asynchrony cannot be solely attributed to endogenous attention effects (Viviani & Aymoz 2001; Bedell et al. 2003; Clifford et al. 2003; Nishida & Johnston 2003; Adams & Mamassian 2004; Aymoz & Viviani 2004; Arnold 2005).

Exogenous attention refers to stimulus properties that involuntarily attract an observer’s attention. It can be difficult to disambiguate these effects from processing time differences, as a change that attracts attention involuntarily may seem to have been processed more rapidly. However, the results of a recent study that examined the effects of stimulus salience might be relevant here (Adams & Mamassian 2004). In this study an operational definition of salience was used – a level of performance during change detection tasks completed before the main experiment. This allowed the experimenters to compare color and stimulus speed changes across a range of relative, operationally defined, saliencies. Temporal order judgments were then used in the main experiment. Thus the experimenters avoided the endogenous attention effects related to the use of "defining" and "report" attributes. It was found that stimulus salience influenced perceptual asynchrony magnitude, with salient stimulus changes appearing to occur earlier than less salient changes. However, when the contrasted stimulus changes were matched in terms of detection ease, color changes still seemed to precede physically coincident direction changes (Adams & Mamassian 2004). Thus, to the extent that equivalent detection rates can be regarded as a control for exogenous attention, the results of the study suggest that exogenous attention cannot provide an adequate explanation for color/motion perceptual asynchrony (Adams & Mamassian 2004).

It is possible that operational defined salience controls are inadequate. It would then be difficult to tease apart the effects of salience and exogenous attention from processing time differences. However, this differentiation may be arbitrary. Strong sensory signals, normally associated with high salience, have been repeatedly tied to faster processing times than weaker signals (Roufs 1963; Williams & Lit 1983; Dzhafarov et al. 1993; Mateeff et al. 1995). For this reason the influence of operationally defined salience variance upon color/motion perceptual asynchrony has been cited as evidence for a link between sensory processing times and perceived timing (Adams & Mamassian 2004).

16.6.1 Section summary

Although both endogenous (Clifford et al. 2003; Enns & Oriet 2004) and exogenous (Adams & Mamassian 2004; also see Paul & Schyns 2003) attention seems to influence perceptual asynchrony magnitude, these effects seem to be additional factors in relation to a more fundamental cause. When both forms of attention are controlled (Adams & Mamassian 2004) color/motion perceptual asynchrony persists.

16.7 General discussion

Although experimental findings demand that the brain time account of perceptual asynchrony be modified, some variation of this perspective still seems like the most viable interpretation. The temporal marker account (Nishida & Johnston, 2002) cannot adequately explain color/motion perceptual asynchrony, because asynchrony magnitude changes when theoretical types of temporal event are held constant (Arnold & Clifford 2002; Bedell et al. 2003; see temporal marker matching discussion). The postdictive analysis account (Moradi & Shimojo 2004) is untenable, as it inaccurately predicts that the best conditions for pairing color and motion should occur when color and direction changes are physically synchronous (see postdictive analysis discussion). The attention switching account is inadequate, as the phenomenon persists when the effects of both endogenous and exogenous attention are controlled (Adams & Mamassian 2004).

The fact that perceptual asynchrony magnitude is variable is entirely consistent with the brain time account. Neural processing times are unlikely to be fixed. They are more likely to require variable periods of time to complete, reflecting the computational demands of specific analyses and the adaptive state of the processing system. Color/motion perceptual asynchrony does not seem to occur when direction changes are not apparent (Clifford et al. 2004; Moradi & Shimojo 2004) and it varies as a function of the angular difference between the contrasted directions of motion (Arnold & Clifford 2002; Bedell et al. 2003). In addition to these observations, the systematic influence of salience on perceptual asynchrony magnitude (Adams & Mamassian 2004) also seems to be indicative of a relationship between the time courses of sensory processing and apparent timing.

16.7.1 Implications

What would it mean if there were a relationship between neural processing times and perceived timing, and why does this proposal excite such debate?

The possibility of a relationship between the time courses of sensory processing and perceptual timing has been criticized on the basis that it would lead to systematic timing errors (Libet et al. 1983; Libet 1985; Eagleman & Sejnowski 2000; Rao et al. 2001; Moradi & Shimojo 2004). According to alternate proposals, timing must be determined by some form of interpretive process that compensates for the variable time courses of sensory processing (Eagleman & Sejnowski 2000; Rao et al. 2001). It is entirely probable that perceptual processing is subject to some form of interpretation (Arnold et al. 2005), but given the number of systematic timing errors that correlate with the known dynamics of visual processing (Arden & Weale 1954; Roufs 1963; Arnold & Clifford 2002; Bedell et al. 2003; Paul & Schyns 2003; Adams & Mamassian 2004; Bedell et al. 2006), it would seem that any interpretive analysis must be subject to the temporal limitations of the neural processing upon which it is based.

The possibility of a relationship between the time courses of sensory processing and apparent timing has also been criticized on more philosophical grounds. It has been argued that the brain time account (Moutoussis & Zeki 1997a,b) treats apparent timing quite differently from other sensory processes (Johnston & Nishida 2001). The argument is that all other sensory properties are encoded in the activity of neurons, whereas brain time suggests that apparent timing might be determined by when analyses are concluded in specific regions of the brain. This point is well made. As I mentioned at the outset, human observers can determine the relative timings of a remarkable variety of visual, auditory, and haptic events (Moutoussis & Zeki 1997a,b; Fujisaki et al. 2004; Vogels 2004; Arnold et al. 2005). Obviously the necessary information is encoded because these relative timing judgments can be made and reported. The question is how?

16.7.2 Integrating independent temporal codes

In the spatial domain, it seems that independent codes are first formed and then summated (Rivest & Cavanagh 1996; Hillis et al. 2002; Alais & Burr 2004). It is conceptually possible that a similar strategy is used for temporal perception. Some form of temporal marker could be attached to neural events that signify *when* the event occurred (Dennett & Kinsbourne 1992; Nishida & Johnston 2002). Of course some form of metric would be required; perhaps this could be provided by an internal clock (Treisman 1984; Buonomano & Karmarkar 2002). If this strategy were used, it would be optimal if the theoretical timing stamps (Dennett & Kinsbourne 1992) or markers (Nishida & Johnston 2002) were attached to neural events at early processing stages in order to minimize the influence of further sensory analyses. One problem with this sort of proposal is that it does not mirror the temporal strategies used by the visual system in other contexts.

16.7.3 Coincidence detection

There is a basic computational strategy used repeatedly in vision – coincidence detection. Neurons that behave like nonlinear coincidence detectors have been implicated in many aspects of visual processing, including motion perception and stereo vision (Reichardt 1961; Barlow et al. 1967; Borst & Egelhaaf 1989; Ohzawa et al. 1990). These neurons can be conceptualized as a logical AND *gate*; they will only respond if they receive two sources of input. This property seems to make them ideally suited for serving as the neural basis for timing perception, as they can encode temporal relationships – for instance a given cell might only increase its firing rate when multiple events occur at the same time (Llinas et al. 2002). Because coincidence detectors are sensitive to the relative timings of sensory input, any change in sensory processing times prior to the process of coincidence detection would change the encoded temporal relationship – an explicit prediction of the brain time account (Moutoussis & Zeki 1997a,b).

As we know that the brain uses coincidence detection as a temporal strategy (Reichardt 1961; Barlow et al. 1967; Borst & Egelhaaf 1989; Ohzawa et al. 1990), it seems entirely reasonable to suggest that a similar strategy might be used to encode relative timing (Llinas et al. 2002). The implication of this is that the visual system would not create different timing codes that must then become integrated. Instead, a process of coincidence detection that monitors multiple sources of input would be used. Given the variety of relative timings that can be judged, the relevant neurophysiological substrates would necessarily be located at brain sites that receive inputs from multiple regions of the brain so that multiple sources of temporal information could be monitored. Numerous findings relating attention to perceived timing suggest that the process is probably gated by attention (Treisman & Gelade 1980; Paul & Schyns 2003; Fujisaki et al. 2006). These suggestions are consistent with functional imaging and clinical observations that implicate a network of structures in timing perception – including the left inferior parietal cortex, the left frontal operculum, the pre-supplementary motor area, and the putamen (Aglioti et al. 2003; Coull 2004).

16.7.4 Coincidence detection predictions

16.7.4.1 Timing errors

The most obvious prediction of a coincidence detection account of perceptual timing is that timing illusions should ensue if processing is delayed prior to the process of coincidence detection. This possibility has attracted considerable attention and many experimental results are consistent with it (Arden & Weale 1954; Roufs 1963; Arnold & Clifford 2002; Bedell et al. 2003; Paul & Schyns 2003; Adams & Mamassian 2004; Bedell et al. 2006).

16.7.4.2 Sensitivity changes

The same point in space can be signaled by multiple cues, by color, motion, and luminance (Rivest & Cavanagh 1996). So too can the same point in time. This fact does not seem to have attracted a great deal of attention.

When redundant sources of information are available, the opportunity exists to combine the two sources of information into an integrated code that is more precise than either of the independent cues. This can be achieved by either summating the two independent sources of information (Rivest & Cavanagh 1996; Hillis et al. 2002; Alais & Burr 2004; Ernst & Bulthoff 2004) or perhaps by a nonlinear process of coincidence detection (Reichardt 1961; Barlow et al. 1967; Borst & Egelhaaf 1989; Ohzawa et al. 1990). The two strategies can predict quite different sensitivity changes.

If cue integration is achieved by summing together independent neural signals, sensitivity for combinations of cues can be predicted on the basis of sensitivity to the independent cues. Following optimal linear summation, perceptual decisions based on the summed signals can be more precise than those based on either of the independent signals. The maximal improvement that can be achieved following linear summation is equivalent to the observer's sensitivity to the independent signals multiplied by $\sqrt{2}$. Such improvements have been observed repeatedly in the spatial domain (Rivest & Cavanagh 1996; Hillis et al. 2002; Alais & Burr 2004; Ernst & Bulthoff 2004).

Much greater improvements might be possible following a nonlinear process of coincidence detection. In this context the different signals are not independent as the coincidence detector must monitor multiple sources of information. Therefore the addition of an extra signal does not introduce additional underlying neural noise, but it does increase signal – a situation that could dramatically improve temporal sensitivity. Accordingly, any situation where temporal sensitivity can be shown to improve by a factor greater than a multiple of $\sqrt{2}$, with the addition of an extra timing cue, would strongly implicate a nonlinear combination process. This possibility could be explored both within vision and across other perceptual modalities.

16.8 Conclusions

Physically synchronous alternations of color and motion direction can appear asynchronous. Although the magnitude of this temporal illusion is variable, a causal relationship between the time courses of sensory processing and apparent relative timing still seems like the most reasonable interpretation of the empirical data. Perhaps the strongest support for the brain time account is provided by the observation that color/motion perceptual asynchrony magnitude varies in a manner that is broadly consistent with the known dynamics of sensory processing (Arnold & Clifford 2002; Bedell et al. 2003; Adams & Mamassian 2004; Clifford et al. 2004).

A close relationship between sensory processing and apparent timing could readily be explained if relative timing were encoded by a computational strategy used repeatedly in vision – coincidence detection. The process of coincidence detection could be gated by attention (Treisman & Gelade 1980; Paul & Schyns 2003; Fujisaki et al. 2006) and the neural substrate would be necessarily situated where multiple sources of temporal information can be monitored. According to this interpretation we should be able to induce timing illusions by selectively slowing perceptual analyses, and we might expect temporal sensitivity to improve in a nonlinear fashion with the addition of extra timing cues.

Acknowledgments

This research was supported by a Discovery Project Grant and Australian Postdoctoral Fellowship awarded to DHA. I am grateful to Signy Wegener and to Tom Wallis for comments on earlier drafts of this manuscript.

References

Adams, W. J., & Mamassian, P. (2004). The effects of task and saliency on latencies for colour and motion processing. *Proc R Soc Lond B Biol Sci* **271**: 139–146.

Aglioti, S. M., Fiorio, M., Forster, B., & Tinazzi, M. (2003). Temporal discrimination of cross-modal and unimodal stimuli in generalized dystonia. *Neurology* **60**: 782–785.

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Curr Biol* **14**: 257–262.

Arden, G. B., & Weale, R. A. (1954). Variations of the latent period of vision. *Proc R Soc Lond B* **142**: 258–267.

Arnold, D. H. (2005). Perceptual pairing of colour and motion. *Vision Res* **45**: 3015–3026.

Arnold, D. H., & Clifford, C. W. G. (2002). Determinants of asynchronous processing in vision. *Proc R Soc Lond B Biol Sci* **269**: 579–583.

Arnold, D. H., Clifford, C. W. G., & Wenderoth, P. (2001). Asynchronous processing in vision: colour leads motion. *Curr Biol* **11**: 594–600.

Arnold, D. H., Johnston, A., & Nishida, S. (2005). Timing sight and sound. *Vision Res* **45**: 1275–1284.

Aymoz, C., & Viviani, P. (2004). Perceptual asynchronies for biological and non-biological visual events. *Vision Res* **44**: 1547–1563.

Bair, W., Cavanaugh, J. R., Smith, M. A., & Movshon, J. A. (2002). The timing of response onset and offset in macaque visual neurons. *J Neurosci* **22**: 3189–3205.

Barlow, H. B., Blakemore, C., & Pettigrew, J. D. (1967). The neural mechanism of binocular depth discrimination. *J Physiol* **193**: 327–342.

Barlow, H. B., & Levick, W. R. (1965). The mechanism of directionally selective units in rabbit's retina. *J Physiol* **178**: 477–504.

Bartels, A., & Zeki, S. (1998). The theory of multistage integration in the visual brain. *Proc R Soc Lond B Biol Sci* **265**: 2327–2332.

Bedell, H. E., Chung, S. T. L., Ogmen, H., & Patel, S. S. (2003). Color and motion: which is the tortoise and which is the hare? *Vision Res* **43**: 2403–2412.

Bedell, H. E., Patel, S. S., Chung, S. T. L., & Ogmen, H. (2006). Perceptual consequences of timing differences within parallel feature-processing systems in human vision. In H. Ogmen & B. G. Breitmeyer (eds.), *The First Half Second: The Microgenesis and Temporal Dynamics of Unconscious and Conscious Visual Processes* (245–258). Cambridge, MA: MIT Press.

Borst, A., & Egelhaaf, M. (1989). Principles of visual motion detection. *Trends Neurosci* **12**: 297–306.

Buonomano, D. V., & Karmarkar, U. R. (2002). How do we tell time? *Neuroscientist* **8**: 42–51.

Clifford, C. W. G., Arnold, D. H., & Pearson, J. (2003). A paradox of temporal perception revealed by a stimulus oscillating in colour and orientation. *Vision Res* **43**: 2245–2253.

Clifford, C. W. G., Spehar, B., & Pearson, J. (2004). Motion transparency promotes synchronous perceptual binding. *Vision Res* **44**: 3073–3080.

Coull, J. T. (2004). fMRI studies of temporal attention: allocating attention within, or towards, time. *Cognitive Brain Research* **21**: 216–226.

Cowey, A., & Heywood, C. A. (1997). Cerebral achromatopsia: colour blindness despite wavelength processing. *Trends Cogn Sci* **1**: 133–139.

Dennett, D. C., & Kinsbourne, M. (1992). Time and the observer: the where and when of consciousness in the brain. *Behav Brain Sci* **15**: 183–247.

Dzhafarov, E. N., Sekuler, R., & Allik, J. (1993). Detection of changes in speed and direction of motion: reaction time analysis. *Perception Psychophys* **54**: 733–750.

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science* **287**: 2036–2038.

Enns, J. T., & Oriet, C. (2004). Perceptual asynchrony: modularity of consciousness or object updating? *Vision Sciences Society Abstracts.*

Ernst, M. O., & Bulthoff, H. H. (2004). Merging the senses into a robust percept. *Trends Cogn Neurosci* **4**: 162–169.

Fujisaki, W., Koene, A., Arnold, D. H., Johnston, A., & Nishida, S. (2006). Visual search for a target changing in synchrony with an auditory signal. *Proc R Soc Lond B Biol Sci* **273**: 865–874.

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nat Neurosci* **7**: 773–778.

Gawne, T. J., Kjaer, T. W., & Richmond, B. J. (1996). Latency: another potential code for feature binding in striate cortex. *J Neurophysiol* **76**: 1356–1360.

Hillis, J. M., Ernst, M. O., Banks, M. S., & Landy, M. S. (2002). Combining sensory information: mandatory fusion within, but not between, senses. *Science* **298**: 1627–1630.

Johnston, A., & Nishida, S. (2001). Time perception: brain time or event time? *Curr Biol* **11**: R427–R430.

Kanai, R., Paffen, C. L. E., Gerbino, W., & Verstraten, F. (2004). Blindness to inconsistent local signals in motion transparency from oscillating dots. *Vision Res* **44**: 2207–2212.

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behav Brain Sci* **8**: 529–566.

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (Readiness-Potential): the unconscious initiation of a freely voluntary act. *Brain* **106**: 623–642.

Libet, B., Wright, E. W., Feinstein, B., & Pearl, D. K. (1979). Subjective referral of the timing for a conscious sensory experience: a functional role for the somatosensory specific projection system in man. *Brain* **102**: 193–224.

Lennie, P. (1998). Single units and visual cortical organization. *Perception* **27**: 1–47.

Livingstone, M. S., & Hubel, D. H. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science* **240**: 740–749.

Llinas, R. R., Leznik, E., & Urbano, F. J. (2002). Temporal binding via cortical coincidence detection of specific and non-specific thalamocortical inputs: a voltage-dependent dye-imaging study in mouse brain slices. *Proc Natl Acad Sci U S A* **99**: 445–449.

Mateeff, S., Dimitrov, G., & Hohnsbein, J. (1995). Temporal thresholds and reaction time to changes in velocity of visual motion. *Vision Res* **35**: 355–363.

Moradi, F., & Shimojo, S. (2004). Perceptual-binding and persistent surface segregation. *Vision Res* **44**: 2885–2899.

Moutoussis, K., & Zeki, S. (1997a). A direct demonstration of perceptual asynchrony in vision. *Proc R Soc Lond B Biol Sci* **264**: 393–399.

Moutoussis, K., & Zeki, S. (1997b). Functional segregation and temporal hierarchy of the visual perceptive systems. *Proc R Soc Lond B Biol Sci* **264**: 1407–1414.

Munk, M. H. J., Nowak, L. G., Girard, P., Chounlamountri, N., & Bullier, J. (1995). Visual latencies in cytochrome oxidase bands of macaque area V2. *Proc Nat Acad Sci USA* **92**: 988–992.

Nishida, S., & Johnston, A. (2002). Marker location not processing latency determines temporal binding of visual attributes. *Curr Biol* **12**: 359–368.

Ogmen, H., Patel, S. S., Bedell, H. E., & Camuz, K. (2004). Differential latencies and the dynamics of the position computation process for moving targets, assessed with the flash-lag effect. *Vision Res* **44**: 2109–2128.

Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1990). Stereoscoptic depth discrimination in the visual cortex: neurons ideally suited as disparity detectors. *Science* **249**: 1037–1041.

Paul, L., & Schyns, P. G. (2003). Attention enhances feature integration. *Vision Res* **43**: 1793–1798.

Qian, N., & Andersen, R. A. (1994). Transparent motion perception as detection of unbalanced motion signals. *II. Physiology. J Neurosci* **14**: 7367–7380.

Qian, N., Andersen, R. A., & Adelson, E. H. (1994a). Transparent motion perception as detection of unbalanced motion signals. *I. Psychophysics. J Neurosci* **14**: 7357–7366.

Qian, N., Andersen, R. A., & Adelson, E. H. (1994b). Transparent motion perception as detection of unbalanced motion signals. *III. Modeling. J Neurosci* **14**: 7381–7392.

Rao, R. P. N., Eagleman, D. M., & Sejnowski, T. J. (2001). Optimal smoothing in visual motion perception. *Neural Comput* **13**: 1243–1253.

Reeves, A., & Sperling, G. (1986). Attention gating in short-term visual memory. *Psychol Rev* **93**: 180–206.

Reichardt, W. (1961). Autocorrelation, a principle for the evaluation of sensory information by the nervous system. In W. A. Rosenblith (ed.), *Sensory Communication* (303–317). Cambridge, MA: MIT Press.

Rivest, J., & Cavanagh, P. (1996). Localizing contours defined by more than one attribute. *Vision Res* **36**: 53–66.

Roufs, J. A. (1963). Perception lag as a function of stimulus luminance. *Vision Res* **3**: 81–91.

Schiller, P. H., & Malpeli, J. G. (1978). Composition of geniculostriate input to superior colliculus of the rhesus monkey. *J Neurophysiol* **41**: 788–797.

Snowden, R. J., Treue, S., Erikson, R. G., & Andersen, R. A. (1991). The response of area MT and V1 neurons to transparent motion. *J Neurosci* **11**: 2768–2785.

Sternberg, S., & Knoll, R. L. (1973). The perception of temporal order: fundamental issues and a general model. In S. Kornblum (ed.), *Attention and Performance IV* (629–685). New York: Academic Press.

Treisman, A. M. (1984). Temporal rhythms and cerebral rhythms. *Ann N Y Acad Sci* **423**: 542–565.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cogn Psychol* **12**(1): 97–136.

van de Grind, W. (2002). Physical, neural and mental timing. *Conscious Cogn* **11**: 241–264.

Viviani, P., & Aymoz, C. (2001). Colour, form and movement are not perceived simultaneously. *Vision Res* **41**: 2909–2918.

Vogels, I. M. (2004). Detection of temporal delays in visual-haptic interfaces. *Human Factors* **46**: 118–134.

Williams, J. M., & Lit, A. (1983). Luminance-dependent visual latency for the Hess effect, the Pulfrich effect, and simple reaction time. *Vision Res* **23**: 171–179.

Zeki, S. (1978). Functional specialization in the visual cortex of the monkey. *Nature* **274**: 423–428.

Zeki, S. (2002). The disunity of consciousness. *Trends Cogn Sci* **13**: 173–196.

Zeki, S., & Bartels, A. (1998). The asynchrony of consciousness. *Proc R Soc London B Biol Sci* **265**: 1583–1585.

Zihl, J., von Cramon, D., Mai, N., & Schmid, C. H. (1983). Selective disturbance of movement vision after bilateral brain damage. *Brain* **106**: 313–340.

17

The time marker account of cross-channel temporal judgments

SHIN'YA NISHIDA AND ALAN JOHNSTON

Summary

The human sensory system, at least in its early stages, consists of multiple channels for different modalities (e.g., vision, audition) and for different attributes in each modality (color, motion). Temporal congruency is a critical factor in the binding of signals across channels, but little is known about what representations and algorithms are used for matching. We first analyze this mechanism from a general theoretical point of view and then address the specific mechanisms underlying the perception of color–motion synchrony and audiovisual simultaneity. We hypothesize that judgments about cross-channel temporal relations are based on the comparison of time markers by a mid-level perceptual process. The time markers are amodal tokens that reference salient, figural features extracted from early-level sensory signals. A temporal marker should reference the time a specific event occurs in the world rather than the time the processing of the event completes in the brain.

17.1 Introduction

The human sensory system has a complex architecture. It consists of multiple parallel channels for different sensory modalities (e.g., vision, audition). The channel for each sensory modality is subdivided into multiple parallel channels, each specialized for processing of different attributes (e.g., color, motion). Furthermore each channel consists of multiple serial processing stages. The transmission and processing of sensory information by neural mechanisms takes time, and the amount of time taken varies significantly across channels. For example, in monkey visual cortex, the latency of stimulus onset evoked response is about 40–100 msec in V1, 50–100 msec in MT, 70–160 msec in V4, and 90–180 msec in IT (Bullier 2001). As a result, sensory signals referring to different aspects of a single object or event are spatially and temporally spread over the brain's neural network. To perceive a coherent world, these signals have to be grouped in a manner that allows referencing to a single object or event. Important physical constraints on the grouping of signals are the coincidence and spatial proximity of the physical attributes from which they arise. However, the means by which the temporal relationships between events are determined remains unclear, particularly when the signals arising from those events reside in separate modalities or modules.

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

In this chapter, we will consider the sensory mechanism that associates and compares temporal patterns of stimulation encoded in separate channels to establish a sense of simultaneity, grouping, and/or binding. We first analyze this mechanism from a general theoretical point of view and then address the specific mechanisms underlying the perception of color–motion synchrony and audiovisual simultaneity. Our general conclusion is that cross-channel temporal comparison is based on the comparison of time markers by a mid-level perceptual process. The term "time marker" is used here to refer to the representations employed in making temporal judgments. There may be a number of different representations depending on the nature of the task. We envisage that time markers for cross-channel comparisons could be amodal tokens, each linked to a specific temporal event. By amodal tokens we mean a representation independent of the modality or nature of the stimulus to which it refers that allows matching of markers arising from dissimilar types of stimulus features. This representation might reference salient, figural features extracted from early-level sensory signals. Note that a temporal marker references the time at which a specific event occurs in the world rather than the time at which the processing of the event is completed in the brain. However, like any other sensory representation, it is not assumed that processing will lead to veridical perception in all situations.

17.2 General consideration

17.2.1 Spatial mechanism

In order to gain insight into the mechanisms of temporal signal association, it is useful to briefly look at the mechanisms of spatial signal association that have been studied more extensively in the field of vision research. The basic assumption of spatial association is that sensory signals occurring closely in space are likely to arise from the same physical object or event.

The mechanisms of spatial association are implemented at multiple levels of visual processing. In early within-channel processing they appear as preattentive mechanisms for grouping and segmentation (Fig. 17.1(a1)). As suggested by the Gestalt laws of grouping, sensory signals detected at separate retinal locations are likely to be bound to a degree that depends upon spatial proximity as well as on continuity, similarity, and so on.

On the other hand, the scheme for cross-channel feature association is considered to be different (Treisman & Gelade 1980; Treisman 1999). Features of the same object/event processed in different modules (Fig. 17.1(a2)) are not grouped by early-level preattentive processing but bound by mid-level processing based mainly on spatial coincidence. Attention plays a critical role in this process.

17.2.2 Spatiotemporal mechanism

Like spatial proximity, spatiotemporal proximity is a cue for signal association. That is, sensory signals coming from slightly different spatial and temporal positions are likely

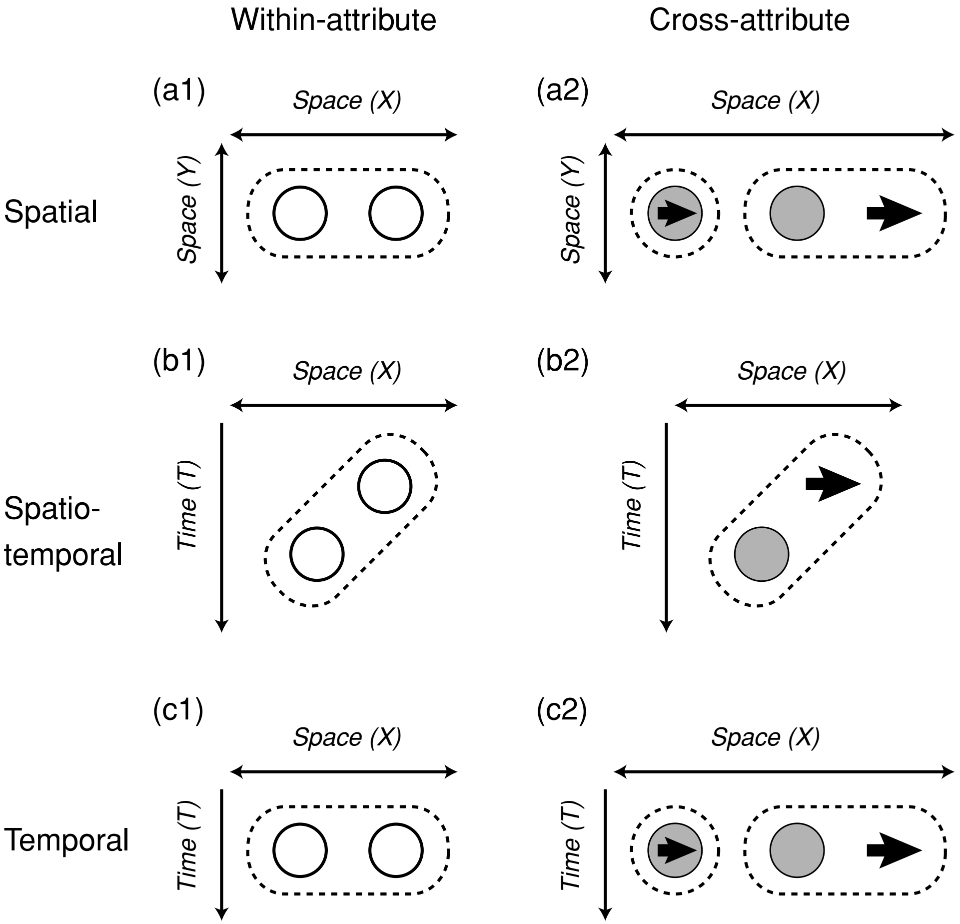


Fig. 17.1 Signal association of various types. (a1) Association between signals of the same attribute type presented at close spatial positions. This within-attribute spatial association is likely to be processed by an early perceptual grouping mechanism. (a2) Association between signals of different attribute types presented at the same (shown on the left) or adjacent (shown on the right) spatial position(s), such as the color and direction of motion of the same object. This cross-attribute spatial association is supposed to be mediated by a mid-level spatial feature binding mechanism. (b1) Association between signals of the same attribute that are presented at adjacent spatiotemporal locations. This within-attribute spatiotemporal association is likely to be facilitated by an early motion mechanism. (b2) Association between signals of different attributes that are presented at adjacent spatiotemporal locations. This cross-attribute spatiotemporal association is supposed to be facilitated by a mid-level feature tracking mechanism. (c1) Association between signals of the same attribute type presented at the same time. This within-attribute temporal association is likely to be sensed by early simultaneity detectors unless the two signals are very much separated in space. (c2) Association between signals of different attribute types presented at the same time. The two signals may be presented at the same spatial location (shown on the left) or at separate locations (shown on the right). We conjecture that this cross-attribute temporal association is mediated by a mid-level temporal feature binding mechanism.

to arise from the same moving object. Indeed, we tend to perceive successive flashes of adjacent lights not as separate objects but as a single moving object. In this respect, the perception of object permanence during motion, especially in the perception of classical long-range apparent motion (Kolers 1972; Braddick 1974), can be regarded as a grouping process based on spatiotemporal proximity.

It has been suggested that the visual system may have two types of motion mechanisms. One is an early short-range mechanism that detects local motion (Braddick 1974; Adelson and Bergen 1985; Johnston et al. 1992), and the other is a mid-level long-range mechanism that links salient elements of the stimulus even across attributes processed in separate modules (Fig. 17.1(b2)) (Braddick 1974). The latter is responsible for interattribute motion (Cavanagh et al. 1989; Lu & Sperling 1995). For example, an element defined only by a color difference from the background may be seen to move into another element defined only by a motion difference, demonstrating that the association cannot be based on the stimulus content. This dichotomy of motion processing could be interpreted to imply that we have two mechanisms for spatiotemporal signal association, as in the case of spatial signal association. Early-level mechanisms are sensitive to within-channel relationships (Fig. 17.1(b1)), whereas mid-level mechanisms are responsible for cross-channel associations (Fig. 17.1(b2)). We think that the mechanisms for temporal association may have a similar duplicity.

17.2.3 Temporal mechanism: preliminary remarks

The main focus of this chapter is the means by which we judge the temporal relationship of events for the perception of simultaneity, synchrony, grouping, and/or binding.

Here we do not separately consider the perceptual mechanisms for binding events based on temporal coincidence and the perceptual mechanisms for judging temporal relations on the order of tens of milliseconds, including simultaneity/synchrony and temporal order judgments. This is because the binding mechanism has to include a process for judging temporal relations, regardless of whether the subjects are explicitly aware of making that judgment or not. On the other hand, when subjects are asked to judge the temporal relations between physically unconnected events, they somehow need to associate the events to make the comparison. We therefore think it likely that the core processing mechanism, or at least the basic processing principle, is common in temporal binding and temporal judgment, although differences in task demand, attentional load, strategy, or criterion applied to judgments might give rise to dissociations in performance (e.g., Clifford et al. 2003).

As is the case in the space and space–time domain, sensory signals occurring closely together in time are likely to arise from the same physical event, and specialized neural processing is necessary to detect the temporal proximity of the signals. The necessity of a detection mechanism, however, is less generally recognized in the case of temporal judgments, presumably because of the conceptual difficulties in appreciating the distinction between physical temporal relations and subjective temporal relations. That is, one can easily make the logical mistake of thinking that physical co-occurrence of “the representation of an event A" and "the representation of an event B" automatically leads to "a representation of the co-occurrence of events A and B" without any additional processing for detecting the co-occurrence of the events (Dennett & Kinsbourne 1992).

There is another challenge specific to temporal judgments. One may assume that the brain uses the timing of sensory neural activity as the basis for temporal judgment. Neural processing, however, is not instantaneous. The time of arrival of neural signals varies for different cortical areas depending on which pathway they belong to and where in the processing hierarchy they reside. In the case of cross-modal association, there is an additional variation in the time of arrival of the physical signal that depends on the transmission media – for example, light or sound. Even in a single cortical area, the neural activity for an instantaneous event is not instantaneous but is spread over time. These issues remind us of the importance of making a distinction between "event time" in the outer world and "brain time" for the processing of the event (Dennett & Kinsbourne 1992; Johnston & Nishida 2001). Ideally, temporal judgments should be based on event time, not brain time. Therefore, neurons should carry "event time" information in such a way that the signal is unaffected by the time required to process it in the brain. Although it may be difficult to completely exclude the influence of some neural delays, especially those arising in very early sensory processing, we do believe that one constraint on the evolution of the brain would be to avoid this type of estimation error where possible.

17.2.4 Early temporal mechanism

To understand how the brain judges the temporal relationships of events, we should clarify the representation and algorithm used for temporal comparison (Marr 1982). The sensory system seems to utilize more than one type of temporal judgment mechanism.

Consider first the situation where the observer has to judge whether two flickering patches of light are in or out of phase. Accurate synchrony judgments are possible at alternation frequencies up to several tens of Hertz when the two patches are adjoining, but performance gradually deteriorates as the size of the gap between the two is increased (e.g., Victor & Conte 2002). This suggests that a spatially localized early visual sensor encodes the temporal relationship of the two patches (Fig. 17.1(c1)). As the spatial gap is increased, the sensor becomes inactive (and presumably synchrony is detected predominantly by a spatially nonlocalized mid-level mechanism as described in the next section). The temporal comparison algorithm used by the early synchrony sensor could be a temporal correlator similar to that proposed by Reichardt (1961) for luminance-based motion detection – there may be sensors sensitive to various amounts of input delay (including zero) across space. It is also possible that some motion detectors themselves operate as synchrony sensors. Directionally balanced motion signals could indicate synchronous events. According to this view, numerous comparators already exist for detecting luminance synchrony between various spatial locations.

Here the input representation for temporal comparison is the relatively raw sensory signals evoked by the stimuli. When the relative neural delay between the compared signals is small and stable, the response of temporal comparators could veridically reflect the time course of the physical events with accuracy on the order of milliseconds (vision, based on motion detectors) or microseconds (audition, based on interaural time difference detectors). Thus, these mechanisms process the time of the event, not the time of the cortical processing of the event. That is, early temporal comparators can detect the temporal synchrony of physical events (e.g., adjoining luminance flicker) regardless of whether the analyses of other properties of the events (e.g., color, shape) complete in other specialized brain mechanisms or not.

17.2.5 Mid-level temporal mechanism

Temporal accuracy is good in cases such as adjacent flicker where early synchrony sensors can operate effectively. However, we are still able to make judgments about the temporal relationships of almost any perceptual events, including those defined by different attributes presumably processed by separate perceptual channels, such as color and motion or light and sound (Fig. 17.1(c2)).

Consider the distinction between explicit and implicit codes. An itinerary provides an explicit route from one city to another, but stray from the route and you are lost. A road map, on the other hand, encodes location – but the route and distance between any pairs of towns can also be extracted. The flexibility of the implicit coding of routes and distances in maps is the principal utility of this type of representation. Rich temporal information of external events is implicitly encoded in many sensory channels in parallel. For judgment of a temporal relation, however, the brain has to form an explicit representation.

In cross-attribute matching, it seems unlikely that the brain prefabricates neural connections and comparators for all possible combinations. Therefore, we suggest that cross-channel comparison may be mediated by mid-level mechanisms that are flexible and less specific to the stimulus that carries the temporal information. This sensory module is likely to be separate from the areas specialized for the processing of given attributes. It could be localized in a specialized cortical area or distributed as a global network (see also 17.3.8). The essential distinction is between a mandatory circuit that delivers up information about temporal properties to decision processes and a visual routine (Ullman 1984) set up as required to perform a perceptual task.

What input representation does the mid-level temporal comparison process use? What are the time markers of events? Alternatives include a timing signal triggered by the stimulus (Libet et al. 1979) or a timing signal associated with the completion of the neural processing of the event in the brain (Zeki 2003). We think the first possibility is more likely, as in the case of the early temporal mechanism, because it prevents temporal judgments from being significantly affected by the time required for the detailed analysis of the event, although the influence of initial neural delays is not excluded. Note, however, we are not proposing that the timing judgments are established as soon as neural responses are evoked by the stimulus. We assume that some time is required for the brain to judge the temporal relation of events, for example, by interrogating the stored dynamic perceptual memory of the events. However, it correctly interprets that the judged temporal relationship refers to past events referenced by the markers (Libet et al. 1979) (see also 17.5).

Representations for cross-channel comparison are likely to be amodal. In audiovisual speech perception, for instance, the pattern of raw neural activity for speech sounds in the auditory cortex and that for mouth movement in the visual cortex have little surface similarity, thus making them hard to compare directly. Ideally, timing signals selected from event-evoked activity should be transformed into common amodal temporal representations (time markers) regardless of the original attributes so that they are suitable for comparison.

A possible candidate for this amodal representation is a saliency map (Itti & Koch 2001; Lu & Sperling 2001). In our usage salience is tightly linked to figure/ground segregation (Lu & Sperling 2001). Salience refers to a graded version of the binary figure/ground assignment. Figural regions have high saliency, whereas ground regions have low saliency. For instance, the crowding effect (He et al. 1996; Pelli et al. 2004) makes a stimulus less salient and brings it down to the state of ground. Salience can be defined independently of the attribute that it describes – it is an amodal descriptor linked to a specific spatial position or temporal instant. One can think of this as a saliency-based intrinsic spatiotemporal map that can be operated on, akin to feature-based intrinsic maps. Salience depends not only on a bottom-up segregation process but also on top-down attentional processes. In many cases, bottom-up processes automatically determine the saliency map, but if conditions are met, one can increase the saliency of a given stimulus by paying attention to it. This enables one to intentionally select time markers from candidates in the stimulus sequence, although this attention-demanding process might suffer severe temporal constraints. The mid-level sensory process then selects, individuates, and registers high-salience features to use for further processing, including cross-channel signal association. Although the amodal saliency map does not carry attribute values by itself, it does have links to the attribute maps.

We conjecture that time markers for cross-channel binding may be associated with salient temporal features (figures in the stimulus stream). We originally conceived of this idea from the analysis of perceptual asynchrony of color and motion (described in detail in the next section). The saliency-based time marker theory is conceptually similar to the idea of a third-order motion system that detects interattribute motion (apparent motion between stimuli defined by different visual attributes, Fig. 17.1(b2)) using the saliency map as input (Lu & Sperling 2001), as well as to the idea that an attribute invariant position tag is used for long-range spatial judgments, including alignment and bisection judgments of Gabor patches (Keeble & Nishida 2001). Figure 17.2 shows the basic structure of the model of cross-channel temporal judgments that we advance.

With regard to the algorithm for mid-level temporal binding, a temporal comparator (cross-correlator) is a promising model. As long as it receives abstracted amodal representations as inputs, a small number of comparators would suffice to process various combinations of stimulus. A line of support for the existence of a cross-channel comparator is recalibration of audiovisual simultaneity (see 17.4.2), which can be ascribed to an adaptation-induced tuning shift of temporal comparators for audiovisual binding, just as

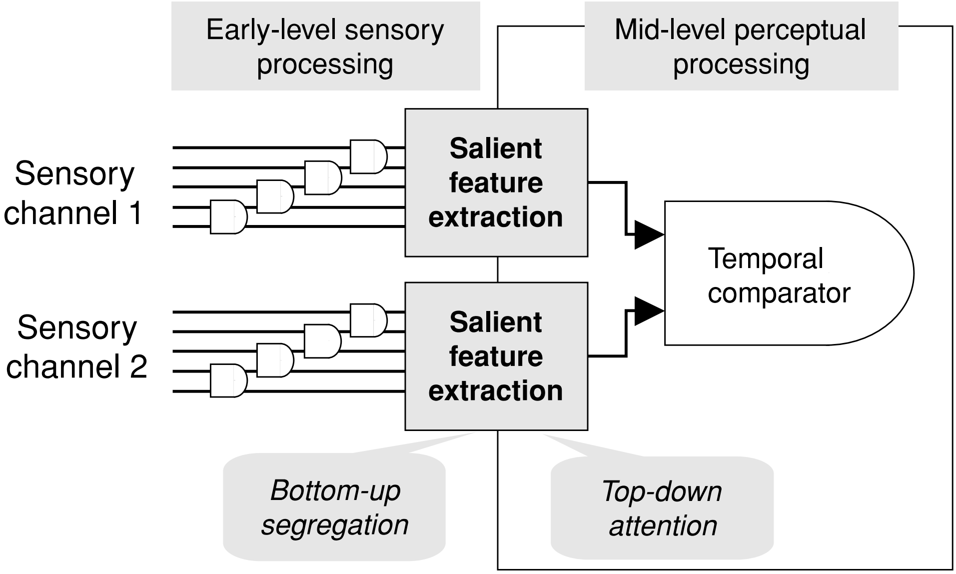


Fig. 17.2 The model of temporal comparison. The cross-channel temporal comparator in the mid-level perceptual system compares the timings of salient features extracted from each sensory channel. The figure also shows early-level within-channel temporal comparators that operate locally.

the motion aftereffect can be ascribed to a tuning shift of motion detectors. Alternatively, a more complex visual routine may be used for mid-level temporal comparison.

How saliency-based cross-channel comparison is implemented neuronally remains an open question. In addition to standard feedforward mechanisms, we do not exclude the possible contribution of neural synchrony often suggested to be involved in temporal judgments and cross-channel binding (e.g., Pöppel 1997; Engel et al. 2001) (see also Section 17.3.8).

17.3 Perceptual asynchrony of color and motion

17.3.1 Phenomenon

Temporal illusions demonstrate that subjective time does not correspond to objective physical time. They also afford great tools for understanding the mechanism of temporal binding and time perception. A good example is color–motion asynchrony (Moutoussis & Zeki 1997a). This is one of the strongest temporal illusions, and it provides profound insight into the neural processes governing subjective time.

In a typical presentation, a green pattern moving upward and a red pattern moving downward are alternated at the rate of 1–2 Hz. At this alternation rate, most observers find it difficult to tell which direction is associated with which color. On the other hand, when the direction change occurs about 100 msec earlier than the color change, the observers reliably bind the two attributes (e.g., red and downward), confidently reporting that the events appear simultaneous. These observations indicate that there is a $\sim$100 msec discrepancy between the point of physical simultaneity and the point of subjective simultaneity.

17.3.2 Processing time hypothesis

Color–motion asynchrony was first reported by Moutoussis and Zeki (1997a). They interpreted this striking phenomenon as reflecting the presence of specialized processing modules for color (V4) and motion (MT/V5) in the visual system. According to their hypothesis, neural activity in each module generates a microconscious experience of the processed attribute. The asynchrony of color and motion results from a processing time difference between color and motion perception. Specifically, processing time for motion is longer than for color by $\sim$100 msec. This hypothesis sounds incompatible with physiological evidence that sensory signals arrive earlier in MT/V5 than in V4. However, if the perceptual asynchrony reflects the times at which processing is completed, not started, and these two times can be dissociated, then this contradiction is resolved.

This processing time hypothesis, however, has a few theoretical problems (Johnston & Nishida 2001), along with several lines of empirical counterevidence (see the following). First, this account has the logical shortcoming of identifying physical co-occurrence of the cortical representation of event A and that of event B with the representation of co-occurrence of events A and B (Dennett & Kinsbourne 1992). Second, it is not clear how the timing mechanism decides when the processing of an event finishes in a specialized processing module. Third, this account assumes a brain time mechanism that is poorly designed in the sense that processing delay is added to event time estimation.

17.3.3 Time marker hypothesis

Many illusions of simultaneity have, like color–motion asynchrony, often been ascribed to differences in neural processing times (Roufs 1963; Stelmach & Herdman 1991; Whitney & Murakami 1998). We accept that relative delays in neural activity may cause some simultaneity illusions, if they occur early enough to affect the relative positions of time markers. For example, there should be a difference in the apparent onset of bright and dim lights as a simple consequence of temporal summation in the retina (Roufs 1963). However, neural delays are not the only cause of temporal misperception, and the comparison process is another source of timing judgment errors. For instance, the comparison process might have a specific tuning bias, as demonstrated in the recalibration of audiovisual simultaneity (Fujisaki et al. 2004). The comparison process might sample the compared stimuli in a temporally asymmetrical fashion, as in the attentional gating phenomenon (Reeves & Sperling 1986). The comparison process might mislabel the temporal order, as in the case of reversed temporal order judgment between crossed hands (Yamamoto & Kitazawa 2001). This might also be the case for reversed temporal order judgments between visual stimuli (Bachmann et al. 2004; Morrone et al. 2005). Finally, the comparison process might use inappropriate features for comparison.

Our hypothesis is that color–motion asynchrony primarily results from matching inappropriate features (Nishida & Johnston 2002). We assume that the visual system assigns temporal markers to salient changes, by means of which temporal localization judgments are made (Fig. 17.2). Here we explicitly make a distinction between the representation used for subjective temporal judgment of an event (the time marker) and the physical time the neural processing of the event takes to complete (and the observer actually sees the event). We also assume time markers are salient (figural) temporal features apparent to the observer in the signal stream. We do not consider putative processing time differences between motion and color as the main factor of color–motion asynchrony, although we do not exclude minor effects of some neural delays that affect the estimation of time marker positions (see Section 17.3.1).

We then ascribe the perceptual asynchrony to the inappropriate matching of color change with position change (motion) rather than with motion direction change. Why does the subject make a matching error of this kind? The point of our argument is that color change and motion direction change are different in temporal structure. To be more specific, color change is a first-order temporal change (first-order temporal derivative of a static attribute that can be defined over two successive points in time), whereas motion direction change is a second-order temporal change (change in the direction of change, a second-order temporal derivative, whose definition requires at least three successive points in time). First-order changes (e.g., changes in luminance, color, position) can be detected in early visual processing by rapid (high temporal resolution) specialized sensors. On the other hand, there is little evidence for the existence of early detectors for second-order changes. For instance, even in the case of acceleration (changes in motion speed), to which some MST neurons are suggested to be sensitive (Kawano et al. 1994), detection by human observers is known to be poor (Gottsdanker 1956; Werkhoven et al. 1992; Simpson 1994). Although the detection of second-order change is possible, it is likely to be subserved by a sluggish (low temporal resolution) mid-level visual routine. Therefore, the comparison between first-order and second-order changes collapses under limiting temporal conditions where second-order changes cannot be used, resulting in a wrong match between available first-order changes (Fig. 17.3).

Our time marker theory was based on several findings. We will consider each of these in turn.

17.3.4 Stimulus temporal structure

According to our hypothesis, it is not a difference in the time required to process each attribute of the stimulus, but the difference in temporal structure, that produces temporal asynchrony. Indeed, we found that second-order temporal changes are apparently delayed relative to first-order temporal changes regardless of the stimulus attributes involved (Nishida & Johnston 2002). That is, just as a motion direction reversal (second-order position change) is apparently delayed relative to a first-order color change, a second-order color change (reversal in the direction of color change) is apparently delayed relative to

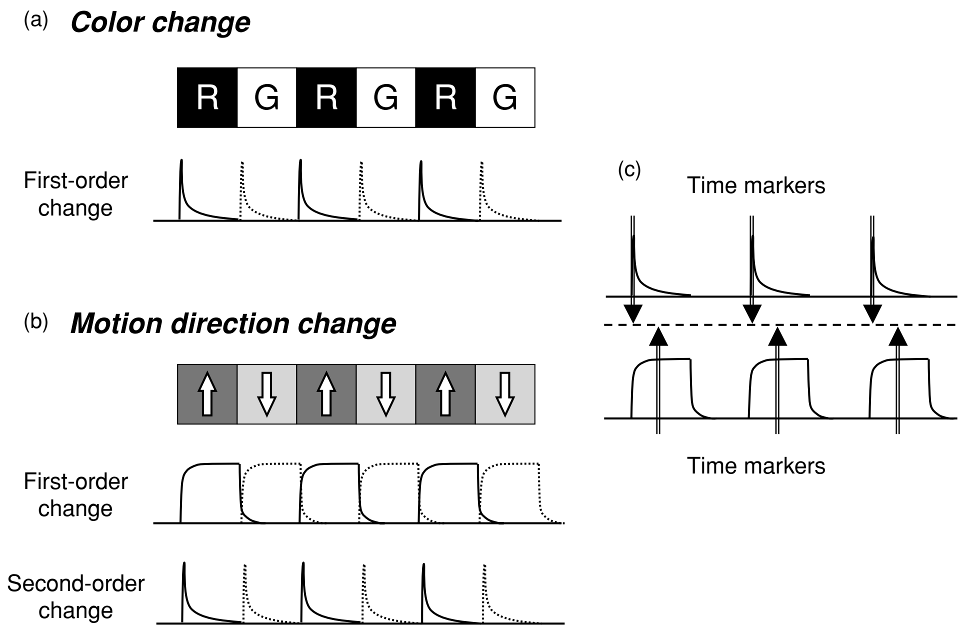


Fig. 17.3 An account of color–motion asynchrony in terms of the matching of inappropriate time markers. (a) A sequence of color change (R: black vs. G: white), and the corresponding first-order change signal. (b) A sequence of motion direction change (upward vs. downward) and the corresponding first-order and second-order change signals. The subject has to match the timing of first-order change of (a) with that of the second-order change of (b). However, at alternation rates of 1–2 Hz, the individuation of second-order change is severely impaired while the detection of first-order change remains intact. In this case, the subject matches the timing of first-order change of (a) with that of the first-order change of (b). The matching process compares time markers assigned to each stimulus change event as shown in (c). Change signals are temporally blurred, but complex neural dynamics including adaptation and mutual inhibition are not considered.

a first-order position change (instantaneous jump). The apparent asynchrony is greatly reduced for the pairing of color and position changes of the same order. Similar results were obtained when color and position changes were compared with luminance changes and with finger/hand position changes. Recently, Arrighi et al. (2005) reported similar findings for an auditory judgment (between frequency modulation and amplitude modulation) and for a cross-modal judgment (between frequency modulation of sound and visual position change).

17.3.5 Alternation rate

Color–motion asynchrony is clearest for repetitive alternation of the stimulus at the rate of 1–2 Hz, regardless of whether the subject explicitly judges apparent synchrony (whether the color change and motion change are in phase) or attribute binding (e.g., whether red is predominantly moving upward or downward). At higher alternation rates, it is difficult to judge the temporal relationship of the two attributes. At slower alternation rates (e.g., 0.5 Hz), simultaneity (in-phase) judgments show little perceptual asynchrony (and it is debatable whether the binding task is meaningful, see 17.3.9). Also, temporal order judgments are veridical (no apparent asynchrony) for comparisons involving a single change of color and a single change in motion direction (Nishida & Johnston 2002; Bedell et al. 2003). These results cannot be easily explained by the processing time hypothesis.

Our interpretation is as follows. For very low alternation rates or single changes, the points of direction change, which are preceded and followed by relatively long stable motion periods, appear to be clearly distinguished as figures in the stream. The time at which there is a change in direction can be marked and properly compared with the time at which the color changes. As the alternation rate increases up to 1–2 Hz, however, the figure–ground relationship is reversed. The points of direction change are subjectively obscured while two directions of motion are still clearly visible. Observers cannot continuously register the timings of motion reversals. We assume that second-order changes are not perceptually delayed but unavailable to the observer under this temporal condition. Instead, it is brief motion periods that are seen as figural features of the sequence. These features are individuated by the mid-level process and compared with the salient features of the color sequence, that is, first-order color changes. For even higher alternation rates, one cannot judge the temporal relationship between color and motion (Arnold 2005; Amano et al. 2007). In this case, a saliency-based temporal binding process does not work because the mid-level process cannot individuate salient features for matching even if they are first-order features. Stated in another way, temporal crowding makes it impossible for the observer to localize each event in the stimulus sequence. The temporal limit of about 3 Hz is consistent with the limit in other vision-related cross-channel temporal binding tasks, such as temporal phase discrimination of color and spatially separate orientation (Holcombe & Cavanagh 2001). This is presumably because these tasks are similarly limited by the temporal resolution of the mid-level visual process.

To account for erroneous use of motion periods (first-order changes) instead of direction reversals (second-order features) at moderately high alternation rates, we look to the contributions of low-level and mid-level mechanisms. Given that sensory signal strength is much higher for first-order changes than for second-order changes (see 17.3.3), it would be natural for a bottom-up stream segregation process to assign first-order changes as figures when they are brief enough to be regarded as instantaneous events. A mid-level process can overcome this low-level decision to some extent, but not rapidly. This is suggested by our finding that temporal order judgments were accurate for a comparison between a single change and the fifth change in rapid alternation regardless of the attribute pairing (Nishida & Johnston 2002). One implication of this finding is that the effect of alternation rate cannot be ascribed to a change in some phase lag of the temporal response to alternation frequency. Another implication is that second-order changes are accessible through temporally modulated attention even at 2 Hz. We suspect that comparing first-order and second-order features requires the setting up of a flexible routine for making the judgment that probably involves temporal modulation of attention and requires a minimum amount of time each time a comparison is attempted. A new stimulus event before this routine has been completed and reset cannot be processed (Raymond et al. 1992). This leads to general cognitive overload, causing the subject to revert to a simpler salience correspondence match between first-order markers. A first-order change (e.g., motion in a particular direction) can be detected directly by rapid sensors and is brief enough to act as a temporal marker. There is no interference to first-order matching from less salient second-order changes at high repetition rates.

17.3.6 Reaction time

Color–motion asynchrony is not accompanied by a behavioral reaction time (RT) difference. We found an apparent $\sim$100 msec advance for color in the color–motion synchrony judgment even when stimuli were chosen such that the RTs to identify stimulus changes were almost identical (Nishida & Johnston 2002). This is in line with a finding by Adams and Mamassian (2004) that responses to motion are much faster than to color when measured in terms of RT, as compared to temporal order judgment. Arrighi et al. (2005) also reported a similar dissociation between perceptual temporal relationships and RT differences for auditory and cross-modal stimulus pairs.

RT can be considered to reflect when task-relevant information becomes available to the observer, plus additional time for motor preparation and execution. This objective time required for a perceptual decision does not have to correspond to the subjective time conveyed by the time marker. It has recently been shown that the time at which the *temporally integrated* population activity of extrastriate cortical areas, including MT/V5, crosses a threshold can quantitatively account for the variation in simple RT to motion onset (Cook & Maunsell 2002; Amano et al. 2006). This result could be interpreted as suggesting that RT is directly correlated with the time at which perceptual processing in a motion-specialized area progresses enough to allow recognition of the stimulus. This corresponds closely to what Moutoussis and Zeki considered as the time of attribute perception. Therefore, the finding that the perceptual asynchrony of color and motion is not accompanied by a corresponding difference in RT provides a strong counterargument to their proposal.

Dissociation between subjective temporal relationships and RT differences has been demonstrated for various stimulus pairs. In many cases, subjective temporal relationships are more veridical than those predicted from RT differences (Tappe et al. 1994; Jaskowski 1996). One reason for this may be that the sensory system somehow compensates for signal delays at the stage of comparison (Fujisaki et al. 2004; Kopinska & Harris 2004), but another reason may be that it uses time markers that are not strongly affected by processing time differences. We think that time markers can be based on a certain temporal index extracted from the time course of stimulus-evoked neural activity in early sensory areas or in more specialized areas. This would be the case for both first-order and second-order changes, although the marker extraction process is more complicated for second-order changes.

It has been suggested that RT reflects the time course of unconscious processing, whereas subjective time reflects that of conscious processing (Libet 1981; Neumann et al. 1993; Tappe et al. 1994). Although we cannot tell whether motor responses in RT experiments are indeed accompanied by conscious awareness, we think this suggestion remains controversial and speculative because it is mainly based on dissociations of subjective temporal judgments and RTs, for which the marker theory provides an alternative account as noted previously.

17.3.7 Motion transparency

A typical color–motion asynchrony display consists of numerous dots that oscillate in synchrony such that the oscillation of the dot pattern and each individual dot is consonant. However, when dots oscillate asynchronously, we see two dot fields moving transparently in the opposite directions (Kanai et al. 2004), and the color–motion asynchrony disappears (Clifford et al. 2004). A similar result was found for motion transparency in which two colors are separately bound to different spatial representations (oppositely moving surfaces) seen simultaneously (Moradi & Shimojo 2004). This finding argues against the processing delay hypothesis because it implies no significant delay between color and motion for the perceptual processes that establish surface representations. On the other hand, this result is compatible with our theory. Our theory predicts that color–motion asynchrony occurs when the attribute binding is dependent on the decision of a mid-level perceptual process about the relative timings of attribute change. In the case of motion transparency displays, an early segmentation process preattentively associates colors and motion directions to the representation of surfaces so that subjects do not have to make temporal judgments and therefore are not prone to misbinding. Cavanagh et al. (2008) demonstrate the same point by using multiple color–motion asynchrony elements and guide the observer's attention by a moving probe.

17.3.8 Cortical response to color–motion asynchrony

Amano et al (2004a,b) studied neural responses in the color–motion asynchrony paradigm using a whole head magnetoencephalogram (MEG). The differences in the time course between MEG responses to color change and motion direction change were significantly reduced when the temporal structure was equated, as expected from the time marker theory. A characteristic MEG response, an enhancement of 30–40 Hz activity in the case of perceptual synchrony (physical asynchrony) between first-order color change and second-order position change (motion), as well as for perceptual synchrony (physical synchrony) between second-order color change and second-order position change was found. This result agrees with previous studies reporting the correlation of gamma band responses with perceptual grouping (Tallon-Baudry et al. 1996; Castelo-Branco et al. 2000), suggesting the contribution of a global neural network in cross-attribute synchrony perception.

17.3.9 Arguments against the time marker account of color–motion asynchrony

Here we address various challenges to the time marker account of perceptual asynchrony.

(i) **The time marker model cannot explain the occurrence of perceptual asynchrony for comparisons between stimulus changes of the same order.** For example, Moutoussis and Zeki (1997b) reported that color change apparently preceded orientation change by 61 msec, although both changes are first order. We have also observed small temporal advantages for color changes relative to position changes, even when these changes were matched in terms of temporal event type (Nishida & Johnston 2002). As noted above, however, there are several factors that can give rise to a discrepancy between subjective simultaneity and physical simultaneity on the order of several tens of milliseconds. Although it is not clear why small asynchronies occur between changes of the same order, it could reflect the effects of attention (Stelmach & Herdman 1991; Paul & Schyns 2003) or stimulus saliency (Adams & Mamassian 2004). It remains open how these factors may influence subjective temporal judgment (see, e.g., McDonald et al. 2005). On the other hand, in the case of color–motion asynchrony (and other combinations of first-order and second-order changes), the size of the illusion can be very large (more than 100 msec). Most importantly, in color–motion asynchrony even physically simultaneous stimuli appear to be perceptually asynchronous. In contrast, in the case of color-orientation binding, whereas the distribution of binding responses is shifted in time due to asymmetric responses to positive and negative asynchronies, physically simultaneous stimuli appear to remain synchronous.

(ii) In disagreement with our finding, Viviani and Aymoz (2001) found that the **apparent time of motion onset was delayed relative to a color change even when they were single isolated changes**. However, the apparent asynchrony is relatively small ($\sim$50 msec) and not robust (Bedell et al. 2003; Adams & Mamassian 2004; Aymoz & Viviani 2004).

(iii) **The amount of asynchrony between first-order and second-order change does not linearly increase as the period of oscillation increases** (Moutoussis & Zeki 1997a; Bedell et al. 2003), although one may expect from our theory that the delay should always be a quarter of the period. However, as the period of oscillation increases, the detection of direction change becomes easier, which would reduce the amount of synchrony (Nishida & Johnston 2002). In addition, even when the first-order change is incorrectly used as a marker, we cannot specify where the marker is localized exactly, because a first-order change lasts for some considerable time. It may be the midpoint of the period if a moving period is short. It may, however, be shifted closer to the motion onset if the period is long, because the subjects know that they should match a motion onset to a color onset (Amano et al. 2007).

(iv) **The time marker theory may account for the judgment of simultaneity, but not for temporal binding**. Color–motion asynchrony occurs both for binding and in-phase (synchrony) judgments at the alternation rate of $\sim$2 Hz. For the slower rate, the two tasks may dissociate (Clifford et al. 2003), but this could be because the subjects can simply ignore obviously asynchronous conditions in the synchrony task while they are forced to choose one of two categories in the binding task. In the latter case, the subjects may be able to see the two phases of color, say in the upward motion interval, and make a quantitative judgment of which color appears to last longer. In other words, subjects could reconfigure the task from a binding task to a duration comparison task, and as a consequence other mechanisms may be recruited. However, as far as we know, when the two attribute changes are perceived to be in synchrony, the binding response is consistent with that perception. Unless this relationship is invalidated, one cannot reject the notion that subjective simultaneity is a basic factor of temporal binding judgments. Recently, Linares and López-Moliner (2006) found that even when the stimulus contained only a single direction change, the use of a binding judgment task, but not a temporal order judgment task, gave rise to an apparent delay of motion relative to color. However, their data showed that when the direction change and color change occurred nearly at the same time, both binding and temporal order judgments were accurate.

(v) **The temporal-phase tuning of the color contingent motion aftereffect also supports asynchrony of color and motion processing** (Arnold et al. 2001). When red changes into green at the middle of the phase of clockwise motion, while the green returns to red at the middle of the phase of anticlockwise motion, each color should be associated with the two motion directions in equal strength if the duration of co-occurrence is the only factor. However, if the response to color is more transient than the response to motion (which is likely considering the order of stimulus change), green will be more strongly associated with the clockwise motion and red more strongly associated with the anticlockwise motion (Johnston & Nishida 2001).

(vi) **The color–motion asynchrony is weakened when the angle of direction change is reduced from 180° (direction reversal) to smaller angles** (Arnold & Clifford 2002; Bedell et al. 2003). Because a direction change is a second-order temporal change regardless of the angle of direction change, the effect of direction angle appears to contradict the time marker theory. We recently investigated this effect systematically (Amano et al. 2007). First, the dependence on the motion direction angle was particularly strong for random-dot stimuli, but our results indicate that this may reflect the introduction of an artifact: motion streaks (Geisler 1999; Burr & Ross 2002) that allow subjects to make a color–orientation synchrony judgment rather than a color–motion synchrony judgment for direction change angles other than 180°. Second, when we used streak-free plaid stimuli, a certain amount of angle dependence remained regardless of whether we asked the observers to judge the apparent binding or synchrony of color and motion direction changes. The degree of direction change also affected reaction times. These findings with plaid stimuli are consistent with the time marker account – the direction change angle can affect the time course of the recruitment of neural responses to the new direction of motion, which will have a consequential effect on the temporal location of salient features in the sequence of motion changes.

17.4 Audiovisual synchrony perception

17.4.1 Time marker theory for audiovisual synchrony perception

To test the general applicability of the time marker theory to cross-channel temporal judgment, we analyzed the mechanism of audiovisual synchrony perception.

If audiovisual synchrony is also detected by a sluggish mid-level process, then the temporal resolution for the detection of audiovisual synchrony is expected to be low. Indeed, for a repetitive pulse train, the upper limit to discriminate a synchronous audiovisual pair from an asynchronous one is only ~4 Hz (Fujisaki & Nishida 2005), regardless of the selection of audiovisual pair (Fujisaki & Nishida 2007). This is much lower than the limits of similar tasks within each modality (vision: ∼25 Hz, audition: ∼80 Hz) that allow a contribution from low-level temporal correlators. For audiovisual pulses of higher frequencies, although the subjects can clearly perceive visual changes and auditory changes, they cannot determine their temporal relationship. This is similar to the difficulty in judging color–motion synchrony at higher temporal frequencies (e.g., 4 Hz).

To be exact, this temporal limit should be evaluated not in terms of the temporal frequency of stimulus modulation but in terms of the density of stimulus features (Fujisaki & Nishida 2007). Even when the stimulus contains low-frequency energy sufficient to support discrimination, synchrony discrimination is not possible for dense random pulses. Furthermore, even when the stimulus itself is dense, synchrony discrimination is possible if it includes sparse salient features (e.g., a red pulse within white pulses, a high-pitch pulse within low-pitch pulses). Feature selection by top-down attention is also useful for audiovisual synchrony discrimination (Fujisaki & Nishida 2008). These results suggest that audiovisual synchrony perception is based on the matching of salient features extracted from each sensory modality.

That feature matching is mediated by a mid-level attentive process rather than an early preattentive process and is further supported by the finding that visual search for a target that changes in synchrony with an auditory stimulus becomes difficult as the number of uncorrelated visual distractors increases – the typical pattern one finds in serial search (Fujisaki et al. 2006). In addition, as noted above, Arrighi et al. (2005) found perceptual asynchrony of audiovisual first-order and second-order changes.

These properties of audiovisual synchrony perception are in agreement with those of spatial, spatiotemporal, and temporal binding between different attributes within vision (Treisman & Gelade 1980; Ashida et al. 2001; Holcombe & Cavanagh 2001), which supports our hypothesis that saliency-based time marker matching is a common principle in mid-level temporal binding.

Our recent data (Fujisaki & Nishida 2009) suggest that the temporal limit of judging synchrony is similarly low for visuotactile pairs. As for audiotactile pairs, the temporal limit is considerably higher than audiovisual and visuotactile pairs, but we also found some properties such as feature invariance and attentional selection that suggest the contribution of a saliency-based matching mechanism.

17.4.2 Recalibration of audiovisual simultaneity

For the veridical estimation of event time, we think it is likely that the brain extracts timing signals from neural responses that are evoked early. However, there is a problem for this strategy in the case of audiovisual binding, because there are time differences in physical transmission (visual signal is faster) and sensory processing (auditory signal is faster) between the two modalities. One strategy the brain might take to overcome this difficulty is to adaptively recalibrate the simultaneity point from daily experience of audiovisual events. Recent studies (Fujisaki et al. 2004; Vroomen et al. 2004) have found novel psychophysical adaptation effects in which exposure to a fixed audiovisual lag for several minutes shifts the center of subjective simultaneity responses in the direction that would reduce the adapted lag. This finding suggests that the brain attempts to adjust subjective simultaneity across different modalities by reducing constant lags between inputs likely to arise from the same physical events. It also indicates that the existence of early temporal discrepancies between channels is not a critical problem for the time marker theory.

17.5 Related theories

Here we discuss the relationships of our theory with other theories of temporal perception that share a view that subjective temporal relationships are a product of the brain's interpretational processes.

The multiple drafts model of Dennett (Dennett 1991; Dennett & Kinsbourne 1992) holds that brain events that discriminate various perceptual contents are distributed in both space and time in the brain, and there is no single, constitutive "stream of consciousness" but rather a parallel stream of conflicting and continuously revised contents. This may sound similar to the microconsciousness theory by Zeki and his colleagues (Moutoussis & Zeki 1997a; Bartels & Zeki 1998; Zeki 2003), but the multiple drafts model makes a strict distinction between time represented (the postmark on the letter, or the time marker of the event) and time of representing (the day the letter arrives, or the time at which the processing of the event is completed). The model explicitly states that whereas the temporal properties of these various brain events are determined, none of these temporal properties determines subjective order. This is identical to our theoretical position. However, the multiple drafts model does not assume specific time markers in the computation of temporal properties across parallel streams, only that the time courses of separately processed attributes may be put in registration at the relative phase of best correspondence (maximum cross-correlation) of the compared contents (which Dennett calls "content-sensitive settling"). We suspect this algorithm is too simple, because it would not distinguish, for instance, synchrony from asynchrony for repetitive changes of color and motion regardless of the magnitude of relative delay.

With regard to the notion of time markers, our theory is much closer to Libet's hypothesis of backwards time referral (Libet et al. 1979), than to the multiple drafts model. Libet found a phenomenon that he considered to imply that a somatosensory stimulation to skin but not a direct electrical stimulation of somatosensory cortex appears to occur at the time of initial cortical response to the stimulus, whereas in either case, establishment of sensory experience of the stimulus (achievement of "neural adequacy") requires cortical activation lasting several hundreds of milliseconds. He explained this by assuming that the brain uses the cortical onset response as the time marker of the event and interprets the processed results of the event to occur at the time referred to by the time marker. There is no paradox in his explanation given that this interpretation is established not at the time of the cortical event onset (which would violate causality) but only after related cortical processing completes. There remain some doubts about the reliability of the experimental findings described by

Libet (Churchland 1981). We do not think there is ample evidence for his hypothesis about the role of an initial cortical response, nor do we accept his argument that RT reflects the time course of unconscious processing and speaks little about the time of establishment of conscious sensory experience (Libet 1981). We however agree with his basic idea that the stimulus-evoked signal is a good time marker of the external event. We think it likely that the cortical processing of an event deepens gradually over time (although one cannot say when the full awareness of the event is established). For example, the RT for the same event (e.g., speech) changes depending on whether the subject has to judge the onset, gender, meaning, or intention of the speech event. However, subjects never make the mistake of thinking that different aspects of the event take place at different times (Dennett 1991). By referring backward to the point indicated by the time marker, we can correctly perceive the temporal structure of the world without being much affected by neural delays.

In agreement with the multiple drafts model and the backwards referral time model, our time marker theory considers subjective event time as a product of retrospective processing. This allows new inputs to affect the interpretation of previous inputs, which Eagleman and Sejnowski (2000) called postdiction. However, our theory has little to do with the specific explanation of the flash-lag effect they proposed (i.e., the flash resets the temporal integration of motion signal). Although the time marker theory, at least in its present form, cannot provide a full account of the flash-lag effect, it predicts that precise temporal judgments are hard because a salient time marker signal is present in flash, while not in continuous motion.

17.6 Conclusion

Here we propose a theory of cross-channel temporal binding. Both theoretical considerations and several lines of empirical evidence support a mid-level matching process based on salient features. Our model, however, remains a conceptual framework rather than a mechanistic explanation. At present, we cannot present a detailed model of how the mid-level perceptual process extracts and compares salient features. Understanding mid-level processing remains a challenge for sensory science.

References

Adams, W. J., & Mamassian, P. (2004). The effects of task and saliency on latencies for colour and motion processing. *Proc R Soc Lond B Biol Sci* **271**: 139–146.

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *J Opt Soc Am A* **2**: 284–299.

Amano, K., Goda, N., Nishida, S., Ejima, Y., Takeda, T., & Ohtani, Y. (2006). Estimation of the timing of human visual perception from magnetoencephalography. *J Neurosci* **26**: 3981–3991.

Amano, K., Johnston, A., & Nishida, S. (2007). Two mechanisms underlying the effect of angle of motion direction change on colour-motion asynchrony. *Vision Res* **47**: 687–705.

Amano, K., Nishida, S., & Takeda, T. (2004a). Enhanced neural responses correlated with perceptual binding of color and motion. *Neurol Clin Neurophysiol* 2004: 48.

Amano, K., Nishida, S., & Takeda, T. (2004b). MEG responses for color-motion asynchrony [Abstract]. *J Vis* **4**: 554a.

Arnold, D. H. (2005). Perceptual pairing of colour and motion. *Vision Res* **45**: 3015–3026.

Arnold, D. H., & Clifford, C. W. (2002). Determinants of asynchronous processing in vision. *Proc R Soc Lond B Biol Sci* **269**: 579–583.

Arnold, D. H., Clifford, C. W., & Wenderoth, P. (2001). Asynchronous processing in vision: color leads motion. *Curr Biol* **11**: 596–600.

Arrighi, R., Alais, D., & Burr, D. (2005). Perceived timing of first- and second-order changes in vision and hearing. *Exp Brain Res* **166**: 445–454.

Ashida, H., Seiffert, A. E., & Osaka, N. (2001). Inefficient visual search for second-order motion. *J Opt Soc Am A Opt Image Sci Vis* **18**: 2255–2266.

Aymoz, C., & Viviani, P. (2004). Perceptual asynchronies for biological and non-biological visual events. *Vision Res* **44**: 1547–1563.

Bachmann, T., Poder, E., & Luiga, I. (2004). Illusory reversal of temporal order: the bias to report a dimmer stimulus as the first. *Vision Res* **44**: 241–246.

Bartels, A., & Zeki, S. (1998). The theory of multistage integration in the visual brain. *Proc R Soc Lond B Biol Sci* **265**: 2327–2332.

Bedell, H. E., Chung, S. T., Ogmen, H., & Patel, S. S. (2003). Color and motion: which is the tortoise and which is the hare? *Vision Res* **43**: 2403–2412.

Braddick, O. (1974). A short-range process in apparent motion. *Vision Res* **14**: 519–527.

Bullier, J. (2001). Integrated model of visual processing. *Brain Res Brain Res Rev* **36**: 96–107.

Burr, D. C., & Ross, J. (2002). Direct evidence that "speedlines" influence motion mechanisms. *J Neurosci* **22**: 8661–8664.

Castelo-Branco, M., Goebel, R., Neuenschwander, S., & Singer, W. (2000). Neural synchrony correlates with surface segregation rules. *Nature* **405**: 685–689.

Cavanagh, P., Arguin, M., & von Grünau, M. (1989). Interattribute apparent motion. *Vision Res* **29**: 1197–1204.

Cavanagh, P., Holcombe, A. O., & Chou, W. (2008). Mobile computation: spatiotemporal integration of the properties of objects in motion. *J Vis* **8**(12): 1, 1–23.

Churchland, P. S. (1981). On the alleged backwards referral of experiences and its relevance to the mind-body problem. *Philos Sci* **48**: 165–181.

Clifford, C. W., Arnold, D. H., & Pearson, J. (2003). A paradox of temporal perception revealed by a stimulus oscillating in colour and orientation. *Vision Res* **43**: 2245–2253.

Clifford, C. W., Spehar, B., & Pearson, J. (2004). Motion transparency promotes synchronous perceptual binding. *Vision Res* **44**: 3073–3080.

Cook, E. P., & Maunsell, J. H. (2002). Dynamics of neuronal responses in macaque MT and VIP during motion detection. *Nat Neurosci* **5**: 985–994.

Dennett, D. C. (1991). *Consciousness Explained.* Boston: Little Brown & Co.

Dennett, D. C., & Kinsbourne, M. (1992). Time and the observer: the where and when of consciousness in the brain. *Behav Brain Sci* **15**: 183–247.

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science* **287**: 2036–2038.

Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top-down processing. *Nat Rev Neurosci* **2**: 704–716.

Fujisaki, W., Koene, A., Arnold, D., Johnston, A., & Nishida, S. (2006). Visual search for a target changing in synchrony with an auditory signal. *Proc R Soc Lond B Biol Sci* **273**: 865–874.

Fujisaki, W., & Nishida, S. (2005). Temporal frequency characteristics of synchrony-asynchrony discrimination of audio-visual signals. *Exp Brain Res* **166**: 455–464.

Fujisaki, W., & Nishida, S. (2007). Feature-based processing of audio-visual synchrony perception revealed by random pulse trains. *Vision Res* **47**: 1075–1093.

Fujisaki, W., & Nishida, S. (2008). Top-down feature-based selection of matching features for audio-visual synchrony discrimination. *Neurosci Lett* **433**: 225–230.

Fujisaki, W., & Nishida, S. (2009). Audio–tactile superiority over visuo–tactile and audio–visual combinations in the temporal resolution of synchrony perception. *Exp Brain Res* **198**: 245–259.

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nat Neurosci* **7**: 773–778.

Geisler, W. S. (1999). Motion streaks provide a spatial code for motion direction. *Nature* **400**: 65–69.

Gottsdanker, R. (1956). The ability of human operators to detect acceleration of target motion. *Psychol Bull* **53**: 477–487.

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature* **383**: 334–337.

Holcombe, A. O., & Cavanagh, P. (2001). Early binding of feature pairs for visual perception. *Nat Neurosci* **4**: 127–128.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nat Rev Neurosci* **2**: 194–203.

Jaskowski, P. (1996). Simple reaction time and perception of temporal order: dissociations and hypotheses. *Percept Mot Skills* **82**: 707–730.

Johnston, A., McCowan, P. W., & Buxton, H. (1992). A computational model of the analysis of some first-order and second-order motion patterns by simple and complex cells. *Proc R Soc Lond B Biol Sci* **250**: 297–306.

Johnston, A., & Nishida, S. (2001). Time perception: brain time or event time? *Curr Biol* **11**: R427–430.

Kanai, R., Paffen, C. L., Gerbino, W., & Verstraten, F. A. (2004). Blindness to inconsistent local signals in motion transparency from oscillating dots. *Vision Res* **44**: 2207–2212.

Kawano, K., Shidara, M., Watanabe, Y., & Yamane, S. (1994). Neural activity in cortical area MST of alert monkey during ocular following responses. *J Neurophysiol* **71**: 2305–2324.

Keeble, D. R., & Nishida, S. (2001). Micropattern orientation and spatial localization. *Vision Res* **41**: 3719–3733.

Kolers, P. A. (1972). *Aspects of Motion Perception*. New York: Pergamon Press.

Kopinska, A., & Harris, L. R. (2004). Simultaneity constancy. *Perception* **33**: 1049–1060.

Libet, B. (1981). The experimental evidence for subjective referral of a sensory experience backwards in time: Reply to P.S. Churchland. *Philos Sci* **48**: 182–197.

Libet, B., Wright, E. W., Jr., Feinstein, B., & Pearl, D. K. (1979). Subjective referral of the timing for a conscious sensory experience: a functional role for the somatosensory specific projection system in man. *Brain* **102**: 193–224.

Linares, D., & Lopez-Moliner, J. (2006). Perceptual asynchrony between color and motion with a single direction change. *J Vis* **6**: 974–981.

Lu, Z. L., & Sperling, G. (1995). Attention-generated apparent motion. *Nature* **377**: 237–239.

Lu, Z. L., & Sperling, G. (2001). Three-systems theory of human visual motion perception: review and update. *J Opt Soc Am A Opt Image Sci Vis* **18**: 2331–2370.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: Freeman.

McDonald, J. J., Teder-Salejarvi, W. A., Russo, F. D., & Hillyard, S. A. (2005). Neural basis of auditory-induced shifts in visual time-order perception. *Nat Neurosci* **8**: 1197–1202.

Moradi, F., & Shimojo, S. (2004). Perceptual-binding and persistent surface segregation. *Vision Res* **44**: 2885–2899.

Morrone, M. C., Ross, J., & Burr, D. (2005). Saccadic eye movements cause compression of time as well as space. *Nat Neurosci* **8**: 950–954.

Moutoussis, K., & Zeki, S. (1997a). A direct demonstration of perceptual asynchrony in vision. *Proc R Soc Lond B Biol Sci* **264**: 393–399.

Moutoussis, K., & Zeki, S. (1997b). Functional segregation and temporal hierarchy of the visual perceptive systems. *Proc R Soc Lond B Biol Sci* **264**: 1407–1414.

Neumann, O., Esselmann, U., & Klotz, W. (1993). Differential effects of visual-spatial attention on response latency and temporal-order judgment. *Psychol Res* **56**: 26–34.

Nishida, S., & Johnston, A. (2002). Marker correspondence, not processing latency, determines temporal binding of visual attributes. *Curr Biol* **12**: 359–368.

Paul, L., & Schyns, P. G. (2003). Attention enhances feature integration. *Vision Res* **43**: 1793–1798.

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: distinguishing feature integration from detection. *J Vis* **4**: 1136–1169.

Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends Cogn Sci* **1**: 56–61.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: an attentional blink? *J Exp Psychol Hum Percept Perform* **18**: 849–860.

Reeves, A., & Sperling, G. (1986). Attention gating in short-term visual memory. *Psychol Rev* **93**: 180–206.

Reichardt, W. (1961). Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In W. A. Rosenblith (ed.), *Sensory Communication*. Cambridge, MA: MIT Press.

Roufs, J. A. (1963). Perception lag as a function of stimulus luminance. *Vision Res* **3**: 81–91.

Simpson, W. A. (1994). Temporal summation of visual motion. *Vision Res* **34**: 2547–2559.

Stelmach, L. B., & Herdman, C. M. (1991). Directed attention and perception of temporal order. *J Exp Psychol Hum Percept Perform* **17**: 539–550.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1996). Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *J Neurosci* **16**: 4240–4249.

Tappe, T., Niepel, M., & Neumann, O. (1994). A dissociation between reaction time to sinusoidal gratings and temporal-order judgment. *Perception* **23**: 335–347.

Treisman, A. M. (1999). Solutions to the binding problem: progress through controversy and convergence. *Neuron* **24**: 105–110, 111–125.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognit Psychol* **12**: 97–136.

Ullman, S. (1984). Visual routines. *Cognition* **18**: 97–159.

Victor, J. D., & Conte, M. M. (2002). Temporal phase discrimination depends critically on separation. *Vision Res* **42**: 2063–2071.

Viviani, P., & Aymoz, C. (2001). Colour, form, and movement are not perceived simultaneously. *Vision Res* **41**: 2909–2918.

Vroomen, J., Keetels, M., de Gelder, B., & Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Brain Res Cogn Brain Res* **22**: 32–35.

Werkhoven, P., Snippe, H. P., & Toet, A. (1992). Visual processing of optic acceleration. *Vision Res* **32**: 2313–2329.

Whitney, D., & Murakami, I. (1998). Latency difference, not spatial extrapolation. *Nat Neurosci* **1**: 656–657.

Yamamoto, S., & Kitazawa, S. (2001). Reversal of subjective temporal order due to arm crossing. *Nat Neurosci* **4**: 759–765.

Zeki, S. (2003). The disunity of consciousness. *Trends Cogn Sci* **7**: 214–218.

18

Simultaneity versus asynchrony of visual motion and luminance changes

MARTIN J. M. LANKHEET AND WIM A. VAN DE GRIND

Summary

Much work has been described comparing relative timing of different features, mostly motion and color or motion and a flash. Here we study the timing relations of pairs of motion stimuli and pairs of motion and flicker or motion and flashes. In a two-alternative forced choice task we measured thresholds for detecting asynchrony, providing estimates for shifts in subjective simultaneity as well as the window of synchronicity.

Windows of synchronicity varied for different combinations of motion direction. Comparing different velocities or different contrast levels revealed large shifts in subjective synchronicity. Contrast effects were much larger for motion reversals than for luminance flicker, indicating a major influence on motion mechanisms. Our results are compatible with the hypothesis of a flexible, high-level brain program for timing analysis. Temporal resolution of this program is limited. Differences in the processing of separate motion characteristics should be taken into account in cross-feature comparisons involving visual motion information. Results for motion reversals versus luminance flashes did not reveal a clear differential shift in time. Large differences within the motion system and the lack of a differential latency between motion reversals and flashes suggest that the flash-lag effect may be largely caused by instant spatial remapping of positional information for moving objects. We show that spatial extrapolation does not necessarily result in overshoot errors when the motion stops.

18.1 Introduction

How long does it take before an elementary physical event like a light flash is consciously perceived? Already in 1868 Donders tried to answer this seemingly simple question by estimating the perceptual latency of flashes from reaction times. He did not find a satisfactory solution to the problem of separating motor latencies from perceptual latencies, nor did anybody else. In his thesis of 1923 Hazelhoff proposed an improved method, which was published a few years later (Hazelhoff & Wiersma 1925). The idea was to use eye movements to time the perceptual flash latency. A vertical dark bar moved horizontally across a screen and was followed by the eyes. The screen had a visible scale above the motion track. In the middle of the track a white bar was briefly flashed, centered on the dark

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

bar. Observers used the scale to determine the perceived position of the moving target at the moment they perceived the flash. They reported a spatial shift Δx in the motion direction, which was interpreted as a latency Δt (with $\Delta t = \Delta x/\text{speed}$). The average perceptual latency for forty subjects was 104 msec for photopic luminances but increased to 250 msec near the absolute threshold for flash detection. Hazelhoff and Wiersma reasoned that the flash was being processed while the eye continued to pursue the dark bar and that the eye's momentary looking direction (the tracked object's position) could be read off as soon as the flash was perceived. The latency of reading the scale position would be irrelevant because it came after the registration of position information.

Fröhlich (1923) showed that the conscious localization of a moving target also has a latency, which ranges from about 40 msec at high to 150 msec at low luminances. A moving bar is first seen at some distance from the edge of a window that it suddenly enters (the Fröhlich effect). Metzger (1932) showed that the sudden (flash-like) appearance of a moving light bar at the edge of a window, as in Fröhlich's experiment, gives a longer latency than the later motion of that same bar. Then he analyzed the phenomenon of Hazelhoff and Wiersma (which we will call the Hazelhoff effect) and argued that a pursuit movement is not necessary to get the effect. Essential is motion of a screen image, such as the screen's scale, across the retina. Because a flash has a longer latency than motion, as his previous experiment showed, the image of the screen will move on while the flash is being processed, so one sees a lagging flash position. Metzger clearly used differential latency as an explanation of the Hazelhoff effect, a hypothesis that is still popular to explain the flash-lag effect (FLE) discovered by Nijhawan (1994). Additional experiments performed by Nijhawan (2001) confirmed Metzger's thesis that the Hazelhoff effect is a species of what we now call the FLE.

Nijhawan (1994, 1997) hypothesized that we see the position of moving stimuli veridically at the perception time of the flash, which requires a positional correction to compensate motion latency. This is an extremely interesting hypothesis, because it would make the speed of visual motion a golden standard for measuring distances and times, like the speed of light in relativity theory. Hazelhoff's original interpretation would then prove to be correct after all. Purely spatial extrapolation implies that the latency difference between flash perception and motion perception might as well be zero. To test this idea, we measured the differential latency between a flash and motion perception as part of a more general project outlined in the following. In the discussion, we will present a purely spatial extrapolation model and show that arguments against the extrapolation hypothesis are based on an incomplete or ill-conceived model.

All time measurements are necessarily based on a process of simultaneity detection, as emphasized a century ago by Einstein (van de Grind 2002). If you cannot ascertain that an event coincides with a certain tick (or group of ticks) of your clock, you cannot measure its time of occurrence. More generally, time measurement requires two ordering operations and a metric, where the ordering operations are simultaneity (equality) and successiveness (inequality). The metric of physical time is given by the ticks of the cesium clock. Experienced time appears to have a different mathematical structure, however, as can be concluded from data reported by Pöppel (2000). The asynchrony threshold, that is the time shift of one relative to another event that just allows us to perceive them as nonsimultaneous, is 5 msec for two auditory clicks, 10 msec for two touches, 20–30 msec for two flashes. However, the threshold for a temporal order judgment (TOJ) is 30–40 msec in all these cases. This means that if A can be perceived to be nonsimultaneous with B ($A \neq B$), we can still find that neither $A < B$ (A precedes B), nor $A > B$. This should warn us not to equate experienced and physical time, and it cautions us not to assume automatically that a TOJ is relevant to understand a simultaneity/asynchrony judgment.

One explanation of the unexpected discrepancy between TOJ and asynchrony data is that the two ordering relations depend on different systems of analysis. In view of this, we will concentrate exclusively on simultaneity versus asynchrony (S/A-) detection, an operation of more general interest. For example, Dennett and Kinsbourne (1992) criticize the common materialistic version of Descartes's dualism, in which it is assumed that everything that enters consciousness should arrive simultaneously at a common finish line, which they call the Cartesian theater. They propose that there are multiple finish lines. We think this is quite plausible. It is an empirical matter to determine how many finish lines there are and what the properties of their S/A operations are. One example of a finish line in action is the integration of simple auditory and visual stimuli in low-level multisensory neurons. According to Stein and Meredith (1993) the decisive factor in getting a strong response of a multisensory neuron in the colliculus of mammals is the temporal overlap of their input signals. Because auditory transduction is faster than visual transduction, one has to present the visual stimulus before the auditory stimulus to get an optimal response of an audiovisual neuron in the colliculus.

There is a direct similarity with psychophysical studies using simple auditory and visual stimuli. These are experienced as optimally simultaneous for a certain auditory lag time, a stimulus onset asynchrony (SOA), which will be called the PSS or point of subjective simultaneity. The probability of experiencing an audiovisual stimulus pair A and B as simultaneous, $p(A \sim B)$, often looks like a bell curve, gradually falling from a maximum simultaneity rating at the PSS toward certainty of asynchrony at shorter and longer SOAs. Simultaneity–probability $p(A \sim B)$ as a function of SOA is the most basic fingerprint of a common "finish line" mechanism for A and B. As soon as $p(A \sim B) = c$, with c some low criterion value (e.g., 25%), the stimulus pair is mostly experienced as asynchronous so these points are called the asynchrony thresholds. One of the two asynchrony threshold-values will be smaller, the other larger than the PSS. The "window of simultaneity" covers all SOAs between the lower and upper asynchrony threshold, and it is often positioned symmetrically around the PSS.

Speech sounds and speech movements of the mouth may be shifted almost 250 msec relative to the PSS before they are experienced as asynchronous, whereas the window of simultaneity for elementary auditory–visual stimulus pairs is only about 80–100 msec. This illustrates that there must be multiple auditory–visual finish lines. One may be in the colliculus to bind elementary audiovisual events in the service of directing eyes and attention.

Another might be found at higher cortical levels to support verbal communication. This suggests that "binding" occurs at various processing levels. Features of the world belonging together because they together constitute causal networks or form a Gestalt are expected to have common finish lines in brain regions specialized in their binding or integration. Incidental connections, such as color and form, might lack a common time-critical finish line. Such features could be bound in working memory through common ranges of local signs, as in a coloring book. This informal idea about binding would be compatible with one interpretation of findings by Moutoussis and Zeki (1997a,b) and Zeki and Moutoussis (1997), suggesting independence of color analysis in V4 and motion analysis in V5 (Bartels & Zeki 1998; Zeki & Bartels 1998a,b,1999; Zeki 2003). However, there are complications. Because it proves possible to measure simultaneity of color and motion stimuli in psychophysical experiments, we must have some neural system for temporal comparison. This common finish line, enabling us to measure S/A for motion and color, contradicts the conclusion of (complete) independence (see Nishida & Johnston 2002).

A temporal comparison of color and motion, as in the cited experiments, has the disadvantage that it is impossible to match the "strength" dimensions. In the luminance domain we know that latency increases for decreasing strength, so by analogy if we find a latency between color and motion it might be caused by nonmatching strengths. To circumvent this matching problem, we decided to initiate S/A detection experiments using feature domains that can be easily matched in strength and have been studied extensively by psychophysical and electrophysiological means: motion and contrast/luminance. The method introduced by Moutoussis and Zeki is convenient and fast, so we decided to use a variant of it. They paired periodically changing motion directions with periodically changing color and varied the relative phase. The method has a drawback that deserves some attention. In the Moutoussis-Zeki experiments the responses in V4 and V5 will show an enforced periodic variation. Simultaneity detection of color and motion direction may, therefore, not only depend on timing of direction and color reversals, but also on the correspondence of temporally integrated half-period responses. If the periodic waves for color and motion have thoroughly different waveforms (as we expect), their optimal matching SOA will include a phase shift component due to waveform differences. Therefore the Moutoussis-Zeki task might not give information on a pure latency difference. Similar reservations were formulated by Nishida and Johnston (2002).

In view of these problems, we deemed it necessary to analyse S/A of zero crossings, that is the simultaneity of the *change events* for each of the features, rather than their simultaneous presence. In addition to the intrinsic interest of this approach to the problem of timing conscious experiences, it also enabled us to evaluate certain explanations of the FLE (flash-lag effect). We studied the S/A profile of a flash and motion in a situation where the spatial and luminance properties were identical. All motion stimuli are random dot patterns, moving within static windows. Flashes are single frames of such patterns. To time changes rather than pairings of features we used a successive two-alternative forced-choice approach, in which one of the alternatives had an objectively or subjectively (see methods) synchronous change of the feature pair. Observers were instructed to choose which of the two successive presentations was the “most synchronous” pair in terms of the direction reversals and/or onsets. To prevent misunderstanding we will use the terms “synchrony” and “window of synchronicity” in relation to our experiments on simultaneity of onsets/offsets and the above terminology for simultaneity of the features themselves, as in the Moutoussis-Zeki experiments.

Among other things, we report that when luminances decrease, motion sensors add latency in addition to and even more than preceding stages. S/A detection is nevertheless possible and must therefore be done across processing levels. This finding contradicts an assumption by Johnston and Nishida (2001) to the effect that temporal analysis only concerns the relative timing of *external* events. The existence of a multilevel temporal analysis also challenges the independence assumption of Zeki and coworkers. We found additional evidence against their independence assumption when we compared the timing of two increasingly different speeds. This leads to a shift of PSS values away from zero, similar to the findings for color and motion by Zeki and coworkers. Following their logic, this would make it necessary to postulate separate consciousnesses for many speed ranges and several luminance levels. We prefer the idea that every common finish line for pairs of features implies the possibility of binding them and that the width of the window of simultaneity or synchronicity indicates how critical timing is to achieve binding. For example, we also report that the window of synchronicity is very narrow for pairs of equal-speed motions in the same direction. Our observers clearly perceive this as causally coupled motions, namely as one big sheet moving behind a screen with two cutout windows. Causal couplings like this one will in general be time critical and thus require a narrow window of simultaneity or synchrony.

18.2 Methods

18.2.1 Procedures

We measured the observer’s sensitivity for detecting asynchrony between two stimulus alternations. Each of the two stimuli consisted of a sparse random dot pattern, shown in a separate window, one on each side of the fixation mark. The stimuli in each window alternated in time between two states. The two states differed in direction or speed of dot motion and/or luminance of the dots. For example, the left window contained dots alternating between leftward and rightward motion and the right window contained a similar pattern alternating between upward and downward motion. In all cases, we used white dots on a dark background. The alternation rate for the two stimuli was the same, but the moments of change of one were shifted relative to those of the other. In the first set of experiments we compared motion direction reversals for different combinations of directions. In a second set of experiments, we compared direction reversals for different speeds of the two moving patterns. In additional experiments, we also compared motion direction reversals to alternations of luminance contrast and brief flashes. The latter experiment is most closely related to classical flash-lag experiments.

The aim was to measure both the two asynchrony thresholds (A preceding B and vice versa) and the PSS. In order to measure a minimal, objective asynchrony threshold, we used a temporal two-alternative forced choice (2AFC). A single trial consisted of an objectively or subjectively synchronous reference and a test with a variable time shift. In situations in which objectively synchronous alternations appeared asynchronous, a subjectively synchronous reference was used (see the following). Test and reference were presented in random order and were separated by a brief interstimulus interval of fifteen frames and an auditory warning signal. The observer indicated, by key presses, which of the two presentations was "most synchronous." Notice that the task did not require the observer to indicate whether or not stimulus alternations were perceived as simultaneous. Results were therefore independent of the observers' subjective internal criterion for simultaneity. Using the 2AFC method yields objective, minimal thresholds. Thresholds were determined using a method of constant stimuli. A suitable range of time shifts was established in pilot experiments. A single experimental session then consisted of six repeats of fifteen settings for the time shift. Sessions were repeated six times, and percentage correct answers were calculated for the thirty-six presentations of each of the fifteen settings.

18.2.2 Stimulus details

Random dot patterns were presented on a Sony 19" multiscan monitor, set to a resolution of 800 × 600 pixels and a refresh rate of 120 Hz. Resolution for the temporal difference between two alternations was given by the frame exposure duration of 8.33 msec. Each random dot pattern consisted of 1000 dots, corresponding to single monitor pixels that were displayed on a dark background (0.5 cd/m^2). By default the luminance of the dots was set to the maximum value (85 cd/m^2). For low-contrast stimuli, the luminance was set to 15 cd/m^2. For flicker stimuli the contrast alternated between these two luminance levels. Except for experiments in which we compared effects for different speeds, speeds were 2 p/f (pixels per frame), which equaled 4.8 deg/sec. The patterns were viewed from a distance of 1.25 m, at which the stimulus windows of 200 × 200 pixels extended 4 deg × 4 deg. The windows were separated by a 2 deg dark gap, in the middle of which a white fixation marker was presented continuously. Observers viewed the stimuli using a chin and forehead rest. All dots were redrawn on every frame of the monitor. In the case of moving dot patterns, the dot positions were shifted by an integer number of monitor pixels on each frame. Dots passing the boundaries of the stimulus window were extinguished and an equal number was regenerated at new random positions along the opposite border. Motion parameters and luminance settings for the dots alternated between two possible states. Unless indicated otherwise, the two states had equal durations. At the standard alternation frequency of 2 Hz (60 frames per cycle) each phase lasted 30 frames (250 msec). A single presentation lasted 2 sec, that is, two complete cycles that contained three alternations. Time shifts between the two stimuli were chosen symmetrically around zero. The value of zero (objective synchrony) was also included in the set.

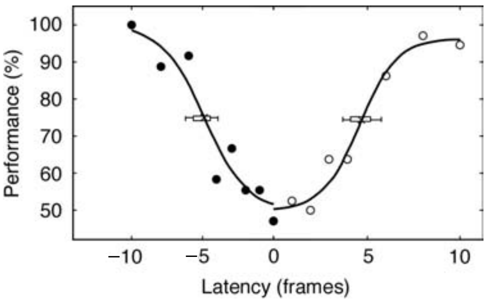


Fig. 18.1 Psychometric curve and fits. Percentage of correct performance was measured for latencies varying from –10 to 10 frames (@120 Hz). Thresholds for detecting asynchrony were obtained by separately fitting a logistic function to positive (open symbols) and negative time shifts (filled symbols). Reliability of the fit and confidence intervals on fitted parameters were determined using Monte Carlo simulations according to the method outlined by Wichmann and Hill (2001a,b). Thresholds were taken at the 75% correct level and confidence intervals at +/–1 standard deviation.

Most experiments yielded symmetrical results for positive and negative time shifts. This places the PSS at zero delay, indicating that internal processing and comparisons did not introduce any systematic differential latencies. In these cases the procedure provides an estimate of p(A∼B) and thus of the temporal window of synchronicity and the asynchrony thresholds. In part of the experiments, synchronous alternations appeared to be processed with a latency difference, leading to an asynchronous percept for the reference stimulus with zero time shift. For example, moving random dot patterns of unequal contrast were generally perceived as asynchronous, even when the physical alternation was perfectly synchronous. In these cases the choice of zero time shift for the reference was inappropriate. Differential latencies for the two stimuli were clearly visible in the pattern of results, in the form of asymmetrical response curves. In such cases we estimated the PSS (latency difference) from pilot experiments and repeated the main experiment with a reference that had a time shift equal to the PSS. Now the reference was perceptually the most synchronous presentation again and observers could perform the task as before, producing symmetric estimated p(A∼B) curves centered at the PSS.

18.2.3 Data analysis

We measured percent correct choices as a function of time shift, both for positive and for negative differences. Figure 18.1 shows an example of data for a single experiment and illustrates the analysis procedure. Open symbols show performance for positive time shifts and solid symbols for negative time shifts. Data for positive and negative time shifts were analyzed separately so as to provide a quantitative measure for asymmetries. Thresholds for detecting asynchrony were calculated by fitting a logistic function to percentage of correct data as a function of latency and determining the latency at which observers performed 75%

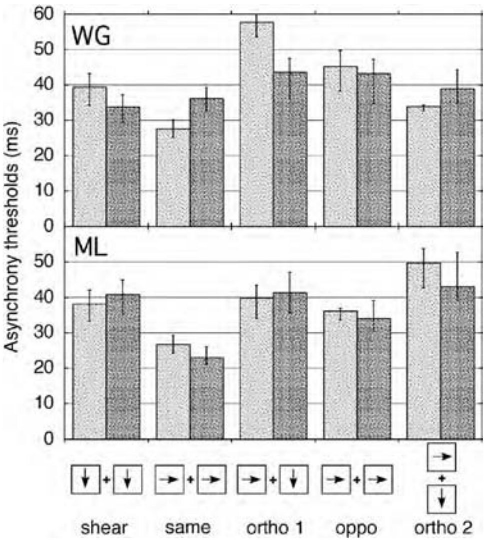


Fig. 18.2 Asynchrony thresholds for motion direction combinations. The left column of each pair represents the threshold for negative time shifts, the right column for positive time shifts in the same experiment. The cartoons underneath a pair of columns illustrate the combination of directions and layout of the two display windows. In each display window, motion directions reversed at a frequency of 2 Hz. Error bars represent +/–1 SD, calculated separately for positive and negative deviations relative to the mean. The experiment labeled “ortho 2” differs in spatial layout, such that information for each stimulus is simultaneously available in both hemispheres. Normally each window is in a different half-field.

correct. Goodness-of-fit and confidence intervals on the estimated threshold were calculated using Monte Carlo simulations (2000 simulations) and bootstrap methods according to the procedure outlined by Wichmann and Hill (2001a,b). We used the 0.159 and 0.841 confidence limits for estimating the reliability of fitted thresholds. Sensitivity for positive and negative deviations from the estimated thresholds was determined separately.

18.3 Results

Figure 18.2 shows data for different combinations of motion directions. In this figure, as well as in all other data figures, results for a single experiment are shown by two columns; the left one represents negative and the right one positive time shifts. Columns of similar height indicate a symmetric window of synchronicity around the reference time shift or PSS. In all experiments shown in Fig. 18.2, the reference time shift was set to zero, in which case symmetry also implies absence of a PSS shift. The cartoons below column-pairs illustrate the corresponding experimental conditions. The smallest window of synchronicity was obtained when the two patterns alternated between the same directions (“same”), that is both left- and right-hand side alternate between left and rightward motion. In this case the window measured about 25 msec (ML) or 32 msec (WG) in each direction. Results for

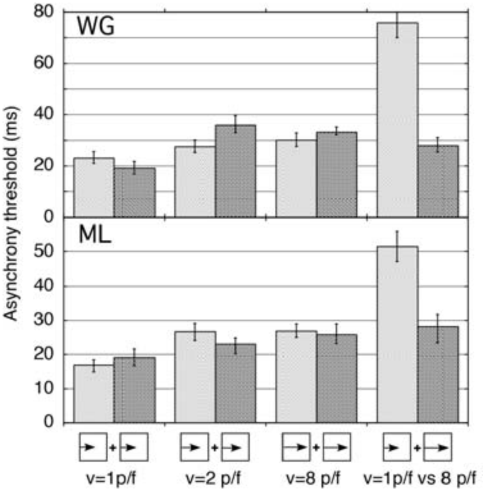


Fig. 18.3 Asynchrony thresholds for different speeds. Format as in Fig. 18.2. In all cases motion directions alternated between leftward and rightward at a frequency of 2 Hz, similar to the "same" condition in Fig. 18.2. Speeds in the left and right window were equal, except for the last column, which shows results for comparing low and high speeds. The large asymmetry for positive and negative time shifts reveals a shift in the point of subjective synchronicity.

other combinations, for example comparing horizontal alternations on the left-hand side to vertical alternations on the right-hand side ("ortho") were considerably higher. Presumably, these different directions of motion lead to similar temporal patterns of activity modulation in the front-end motion system. Yet, judgments of synchrony differ. Results may even differ for directional reversals for horizontal motion. Patterns moving in opposite ("oppo") directions in the left and right window tend to have a larger window of synchronicity than patterns moving in the same direction. From these experiments, we conclude that judgments of synchrony may even differ for stimuli that evoke similar patterns of activity in the front-end visual system. It is not the time course of low-level responses, but additional constraints at some higher level, that determine the width of the window.

In all, except for the last condition visual information from the two windows is initially processed in different hemispheres. Comparing the time course of modulations thus requires information to cross the corpus callosum. To check whether this affects the result, we also performed the experiment as shown on the right-hand side in Fig. 18.2. Here, the two windows are displayed on the vertical meridian, and hence information from both windows ends up in both hemispheres. The result shows that this does not lower asynchrony thresholds.

Figure 18.3 shows results for different speeds. In all cases motion directions alternated between leftward and rightward, similar to the "same" condition in Fig. 18.1. Except for the last column, speeds for the left and right window were equal. For these conditions results were symmetrical and centered around zero. If, however, the task involved comparisons of direction reversals for a high and a low velocity, the result became strongly asymmetrical,

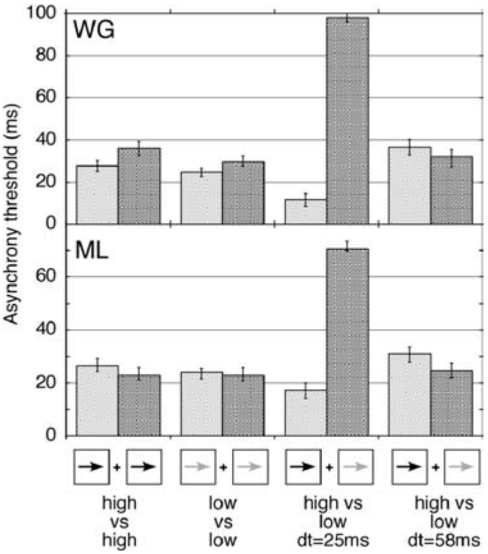


Fig. 18.4 Asynchrony thresholds for different contrasts. In all cases the patterns moved at a speed of 2 p/f, reversing directions at a frequency of 2Hz. Format as in Fig. 18.2. High contrasts corresponded to a luminance of 85 cd/m^2 and low contrast to a luminance of 15 cd/m^2 Contrast differences between the stimuli in the left and right window ("high vs. low") lead to strongly asymmetrical results, indicating large latency differences. These latency differences were partially compensated for by introducing a time shift in the reference stimulus, as indicated by the label "dt = 25 msec." A difference of seven frames (dt = 58 msec) in the reference fully compensated the latency.

revealing different latencies for different velocities. Two effects may play a role here. First, temporal properties for different motion detectors may differ. If higher velocities are detected over shorter time intervals, signals will have a shorter latency for high than for low speeds. Secondly, introducing a velocity difference might rule out the involvement of integrative motion analysis and thus raise asynchrony thresholds. The important conclusion is that a mere change of speed drastically affects temporal judgments.

Besides speed of motion, several other stimulus parameters are likely to influence temporal judgments. It is well known that response latencies in the visual motion system also vary with luminance and contrast. In the next experiment we compared the effect of luminance contrast on asynchrony thresholds (Fig. 18.4). The experiment is similar to that in Fig. 18.3, but we changed luminance rather than speed of motion. The first two conditions show results for two high and for two low contrast stimuli. Although a reduction in contrast is known to introduce a considerable response delay in V5 (MT) motion signals, it has little effect on asynchrony thresholds. Thus, temporal resolution of activity modulations does not seem to limit performance. However, if we compare high versus low contrast (third condition in Fig. 18.4) the result is strongly asymmetrical. A first guess at the stimulus time shift required to compensate for increased latency (25 msec) of the low contrast stimulus was clearly insufficient. Based on data from this experiment, we calculated the PSS and ran the same experiment again using the measured PSS shift as time shift in the reference stimulus.

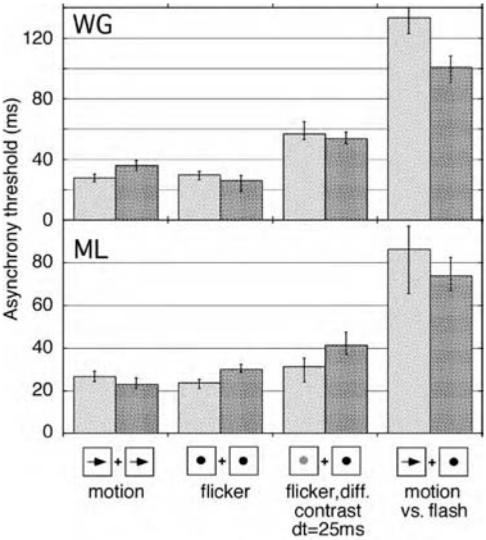


Fig. 18.5 Asynchrony thresholds for motion reversals, luminance flicker of equal and of different contrast, and motion reversal versus flashes. In the standard flicker experiment (second condition) luminance values alternated between the background value and the maximum value (85 cd/m^2). In the third condition the luminance level for one of the two stimuli was reduced to 15 cd/m^2. In flash experiments (rightmost condition) the task was to detect asynchrony in motion reversals and one frame flashes around the direction reversal times. This proved impossible at the standard frequency of 2 Hz. Therefore alternation frequency was reduced to 1 Hz, and total stimulus duration was increased to 4 sec. Asynchrony for flicker at different contrasts was measured using a reference that was shifted three frames (25 msec) to compensate for latency differences.

The result, shown on the right-hand side in Fig. 18.4, was very similar to those of the first two conditions. Contrast clearly has a huge effect on the PSS, a shift of about 60 msec, but it does not affect the window of synchronicity. Even if contrast affects the time course of motion responses, it is not a limiting factor for the temporal comparator.

The differential response latency of about 60 msec presumably results partly from changes in retinal dynamics. If that were the only factor, we would expect similar shifts for a luminance flicker task. If contrast effects are smaller for flicker, this would indicate additional contrast effects on the motion sensors. To find out, we compared the motion task to different flicker tasks (Fig. 18.5). In these flicker tasks the random dot patterns were stationary and luminance levels alternated between the background level and either a high or a low luminance level. For luminance flicker of equal contrast in the two windows (labeled “flicker” in Fig. 18.5) we found results similar to those for equal motion stimuli. For unequal contrast flicker (third condition in Fig. 18.5) we found a modest shift in latency, and if we shift the high-contrast stimulus by 25 msec the effect is almost perfectly compensated. The conclusion must be that contrast effects on motion asynchrony result, to a large extent, from changes within the motion system rather than at earlier (retinal) levels. This is also in line with the size of the Pulfrich effect. For a tenfold decrease in luminance one generally finds a differential delay of about 15–20 msec.

All experiments so far concerned temporal comparisons within a single system, either motion or low-level luminance analysis. Next, we attempted to go one step further and compare motion direction reversals to luminance and color flicker. Notice that the stimuli in these experiments are similar to those of Moutoussis and Zeki (1997a,b), but the observer's task was clearly different. In their experiment observers were asked to judge the simultaneous presence of two features, whereas in our experiments the task required a judgment on synchronicity of transitions. Much to our surprise, this task proved impossible. At the frequency of 2 Hz we were unable to consistently distinguish synchronous from asynchronous presentations. This was true for motion versus color flicker and for motion versus luminance flicker. Reducing the frequency to 1 Hz made little difference. Thus, observers may be able to indicate the best correspondence between features while being unable to detect the asynchrony of the transitions.

Finally, we addressed the question of asynchrony detection between a flash and motion direction reversals related to the FLE. We used a single brief flash (one single frame, luminance 85 cd/m^2) of a random dot pattern in one window and compared it to motion direction-alternations in the other window. This decouples the location information of flashes and motion and focuses on temporal effects. At a frequency of 1 Hz and a duration of 4 sec (effectively three flash motion reversal combinations) one can do the task, albeit with great difficulty. Results are shown in the rightmost column of Fig. 18.5. Asynchrony thresholds were very high, between about 70 and 130 msec, indicating very wide windows of synchronicity. Furthermore, the result did not reveal the strong asymmetry that one might have expected. Both positive and negative asynchrony thresholds were raised considerably, without a large shift in PSS. These findings lead us to propose that temporal delays are less important in explaining the flash-lag effect because processing delays would show up as shifts in the PSS.

18.4 Discussion

Our experiments were designed to find and characterize common finish lines for feature pairs in the motion and luminance domains. A long-term aim is to better understand the consequences of different processing latencies for timing of conscious experiences and for feature binding in vision. Here we used a 2AFC method optimised to detect S/A of changes in motion direction or luminance (flicker and flashes). The method allowed us to measure PSS values (differential latencies) and asynchrony thresholds for various pairs of features. Results are relevant to characterize the properties of combination and comparison mechanisms at the common finish lines.

18.4.1 Temporal shifts, windows of synchronicity, and binding

The window of synchronicity is highly relevant for binding of different perceptual features. Outside this window, features are segregated in time and hence oppose binding into a single object. Our finding is that windows of synchronicity are relatively wide, even for comparisons of modulations within the motion analysis system. Relative latency differences up to about plus and minus 30–40 msec pass unnoticed, indicating that the time analysis has a low temporal resolution. This is surprising, considering the high temporal frequency sensitivity of low-level feature detectors. Directionally selective units in macaque area MT or cat area PMLS, which can be considered second stages for motion processing, respond reliably to direction reversals at rates up to 120 Hz (Borghuis et al. 2003; Perge et al. 2004; Vajda et al. 2004). Obviously, such responses are too fast to encode in spike rate modulations and must be based on accurate spike timing. Highly reproducible spike timings for motion responses have also been reported by Bair and Koch (1996). Given the low temporal resolution of time difference analysis we must assume that spike timing information is no longer available at the levels studied in this chapter. Instead, detection of temporal delays seems to involve considerable temporal integration.

It is not directly clear why different direction combinations should yield different results. It seems reasonable to assume that reversals in horizontal and vertical directions generate similar temporal activity profiles in low-level motion sensors. Yet, it matters which reversals are being compared. The time comparator could of course be more efficient for specific direction combinations, but there are no other indications that this might be the case. An alternative explanation is that some combinations pass additional stages of processing before temporal comparison. Neurons in area MST, for example, globally integrate local motion responses from area MT, in large receptive fields sensitive to different flow patterns. Integrating responses from the two windows, and across a large population of similarly tuned MT cells, may greatly reduce the noise and improve timing. In this view, the “same” condition would give lowest asynchrony thresholds because it best matches the tuning properties of neurons at the level of area MST.

By comparing different velocities and contrasts we focused on differences in temporal processing within the motion system. These differences were substantial: comparing different velocities may raise the asynchrony thresholds by about 30–40 msec. Contrast differences cause even larger shifts in asynchrony thresholds. Such contrast-induced shifts partly resulted from low-level, presumably retinal, effects. The fact, however, that flicker comparison, with the same contrast differences, yielded only a minor shift (about 25 msec) shows that contrast effects primarily result from changes within the motion system. These changes are therefore highly relevant for any cross-modal comparison involving motion signals. The effects we measured for different motion stimuli are as large as previously reported cross-modal effects, indicating that such effects might be stimulus specific rather than system properties. Choosing different motion parameters in cross-modal comparisons will greatly influence the results.

The most important property to take into account concerns the fact that it takes time to detect motion. This is easily appreciated for motion sensors of the bilocal detector type, similar to a Reichardt detector. In such a model a moving pattern has to travel the distance between two subfields before motion is detected. Depending on the velocity preference, the distance corresponds to a minimal time delay. Estimates of preferred step size and delay tuning in human motion detection showed that tuning for delay varied with speed (van Doorn & Koenderink 1982; Koenderink et al. 1985; van de Grind et al. 1986). For low speeds one finds long delays, whereas for high speeds delays are relatively small. It has been shown (van den Berg and van de Grind 1989) that differences in reaction time data for motion onsets can also be explained along these lines.

The present results make the idea of separate microconsciousnesses per feature unattractive. Our results for motion stimulus pairs of unequal contrast are similar to those reported by Moutoussis and Zeki (1997a,b) for color and motion, so one would have to postulate a microconsciousness for each of a number of levels of contrast. The same conclusion would follow for different speeds of motion. It is more parsimonious to assume that the visual system has one or more temporal analysis modules that can analyze both front-end (as suggested by Nishida & Johnston 2002) and feature-specific stages of vision. Binding might then be the default for responses to stimuli that cannot be temporally segregated. Perhaps features get bound unless the temporal analyzer signals deviations from synchronicity. Windows of synchronicity are, however, fairly wide. This should take care of most latency differences occurring for different features belonging to the same object.

18.4.2 Temporal and spatial factors in the flash-lag effect

There is an ongoing debate on the explanation of the flash-lag effect. Nijhawan's original proposal for extrapolation of positional information from moving objects has been challenged by numerous other studies. A critical piece of evidence against "motion extrapolation" comes from flash-terminated motion sequences or sequences changing direction at the flash. Most studies agree that the flash-lag effect is much weaker or absent altogether in flash terminated experiments. In direction reversal experiments the effect is correlated with motion after the flash (Whitney & Murakami 1998; Eagleman & Sejnowski 2000). At first sight this appears to be direct evidence against temporal extrapolation. Our data, however, provide support for a spatial rather than temporal explanation of the FLE. We measured broad windows of synchronicity, suggesting a sloppy temporal comparison. We did not find the clear shift in PSS that one would expect if flashes and motion would have different latencies. Moreover, temporal processing for motion very much depended on the speed of motion. This would be a serious problem for proper timing of sensory-motor interactions in a dynamic world, unless latencies are properly compensated for. Spatial extrapolation as proposed by Nijhawan is completely in line with all of these findings and it seems worthwhile to critically challenge the counter evidence.

To defuse the evidence against motion extrapolation we propose a mechanism that prevents updating of position information of moving stimuli as soon as the motion stops. We assume that the absence of a flash-lag in flash-terminated experiments shows that the last part of the motion trajectory does not get extrapolated. Figure 18.6 shows a schematized mechanism that could implement this idea (Lankheet & van de Grind in preparation). The network is built up from so-called bilocal motion detectors in which a delayed input and spatially shifted, nondelayed, input are correlated to produce a motion signal output. This output is (possibly via additional interneurons) also used to spatially remap the position

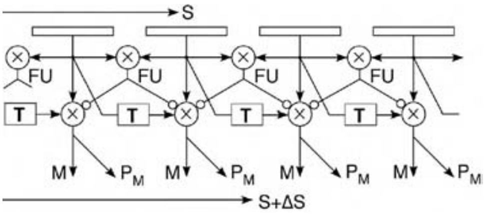


Fig. 18.6 A simplified network model for instant spatial remapping of motion signals, consisting of a chain of bilocal motion sensors. The top layer represents input units, feeding into two different layers of correlation units. Motion is detected by correlating a delayed (indicated by T) input to a spatially shifted, nondelayed input. The motion signal is used to update positional information, spatially compensating for the response latency. Global flicker or uniform light is detected by a correlator (FU unit) combining the same two inputs without a time delay. The FU unit inhibits nearby motion detectors to prevent motion-position remapping when the motion stops.

information. We assume that the mapping of motion sensor outputs onto an internal functional position map is shifted a distance ΔS in the direction of the signaled motion. This implements Nijhawan's extrapolation idea.

In addition, our network model also comprises inhibition from flicker-and-uniformity (FU) detectors. These detectors combine the same two inputs as motion sensors, but without the extra delay. FU outputs spread to a region of motion sensors (only neighboring connections are shown in Fig. 18.6) and inhibit the motion sensor outputs in that region. Such an inhibition is required in bilocal motion sensors to suppress a response to global uniform or flickering, nonmoving stimuli. However, the output of the FU detector will also terminate extrapolation as soon as the motion stops. In the diagram in Fig. 18.6, this is implemented by inhibiting the motion outputs on both sides, but the spread of this inhibition should be tuned to the value of ΔS. This effectively abolishes the spatial remapping of motion information at the point where the motion stops. It should furthermore be noticed that we've drawn only part of the network in Fig. 18.6, namely one string of motion sensors. Similar strings exist for interleaved and/or staggered receptive fields in the input layer. The spread of inhibition therefore is not limited in distance to an integer number of span units of the corresponding motion sensors but is a free (quasi-continuous) parameter. An attractive property of the model in Fig. 18.6 is that it automatically sharpens the borders for moving stimuli in the position map and thus reduces motion smear. Once a border has passed, receptive field pairs of motion sensors are covered by the same unchanging input, so the FU detector is activated. This actively suppresses motion signals and further spatial remapping, limiting the representation to just the contour.

This **gated spatial remapping** principle can explain the FLE as envisioned by Nijhawan (1994, 1997) and is compatible with the findings obtained with flash-terminated displays. We assume that ΔS and the FU spread are tuned during ontogenesis, but to normal daylight circumstances. As soon as the front-end receptive fields have to respond to lower luminances, an extra and uncompensated delay is introduced. This means that the perceived motion position starts to lag the actual motion position, which will decrease the FLE. If the extra front-end delay gives a larger apparent position shift than ΔS, the sign of the FLE can even reverse. These predictions are supported by reports in the literature (Purushothaman et al. 1998; Krekelberg & Lappe 1999). The elegant part of our explanation is that researchers who have vigorously attacked each other all appear to be partly right; an integrative theory along the lines just sketched appears possible. Gated spatial remapping implies that motion, as in relativity theory, can be used as a space and time reference in vision.

References

Bair, W., & Koch, C. (1996). Temporal precision of spike trains in extrastriate cortex of the behaving macaque monkey. *Neural Comput* **8**: 1185–1202.

Bartels, A., & Zeki, S. (1998). The theory of multistage integration in the visual brain. *Proc R Soc Lond B* **265**: 2327–2332.

Borghuis, B. G., Perge, J. A., Vajda, I., van Wezel, R. J., van de Grind, W. A., & Lankheet, M. J. (2003). The motion reverse correlation (MRC) method: a linear systems approach in the motion domain. *J Neurosci Methods* **123**(2): 153–166.

Dennett, D. C., & Kinsbourne, M. (1992). Time and the observer. The where and when of consciousness in the brain. *Behav Brain Sci* **15**: 183–247.

Donders, F. C. (1868). On the speed of mental processes. *Reproduced in Acta Psychologica* **30**: 412–431.

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science* **287**: 2036–2038.

Fröhlich, F. W. (1923). Über die Messung der Empfindungszeit. *Zeitschrift für Sinnesphysiologie* **54**: 58–78.

Hazelhoff, F. F., & Wiersma, H. (1925). Die Wahrnehmingszeit. *Zeitschrift für Psychologie* **96**: 171–188, and **97**: 174–190.

Johnston, A., & Nishida, S. (2001). Time perception: brain time or event time? *Curr Biol* **11**: R427–R430.

Koenderink, J. J., van Doorn, A. J., & van de Grind, W. A. (1985). Spatial and temporal parameters of motion detection in the peripheral visual field. *J Opt Soc Am A* **2**: 252–259.

Krekelberg, B., & Lappe, M. (1999). Temporal recruitment along the trajectory of moving objects and the perception of position. *Vision Res* **39**: 2669–2679.

Metzger, W. (1932). Versuch einer gemeinsamen Theorie der Phänomene Fröhlichs und Hazelhoffs und Kritik ihrer Verfahren zur Messung der Empfindungszeit. *Psychologische Forschung* **16**: 176–200.

Moutoussis, K., & Zeki, S. (1997a). A direct demonstration of perceptual asynchrony in vision. *Proc R Soc Lond B Biol Sci* **264**(1380): 393–399.

Moutoussis, K., & Zeki, S. (1997b). Functional segregation and temporal hierarchy of the visual perceptive systems. *Proc R Soc Lond B Biol Sci* **264**(1387): 1407–1414.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature* **370**: 256–257.

Nijhawan, R. (1997). Visual decomposition of colour through motion extrapolation. *Nature* **386**: 66–69.

Nijhawan, R. (2001). The flash-lag phenomenon: object and eye movements. *Perception* **30**: 263–282.

Nishida, S., & Johnston, A. (2002). Marker correspondence, not processing latency, determines temporal binding of visual attributes. *Curr Biol* **12**: 359–368.

Perge, J. A., Borghuis, B. G., Bours, R. J., Lankheet, M. J., & van Wezel, R. J. (2005). Temporal dynamics of direction tuning in motion sensitive macaque area MT. *J Neurophysiol* **93**: 2104–2116.

Pöppel, E. (2000). *Grenzen des Bewusstseins (Limits of Consciousness).* Frankfurt a.M. & Leipzig: Insel Verlag.

Purushothaman, G., Patel, S. S., Bedell, H. E., & Öğmen, H. (1998). Moving ahead through differential visual latency. *Nature* **396**: 424.

Stein, B. E., & Meredith, M. A. (1993). *The Merging of the Senses*. Cambridge, MA: Bradford Book, MIT Press.

Vajda, I., Lankheet, M. J., Borghuis, B. G., & van de Grind, W. A. (2004). Dynamics of directional selectivity in area 18 and PMLS of the cat. *Cereb Cortex* **14**(7): 759–767.

van de Grind, W. A. (2002). Physical, neural, and mental timing. *Conscious Cogn* **11**: 241–264.

van de Grind, W. A., Koenderink, J. J., & van Doorn, A. J. (1986). The distribution of human motion detector properties in the monocular visual field. *Vision Res* **26**: 797–810.

van den Berg, A. V., & van de Grind, W. A. (1989). Reaction times to motion onset and motion detection thresholds reflect the properties of bilocal motion detectors. *Vision Res* **29**(9): 1261–1266.

van Doorn, A. J., & Koenderink, J. J. (1982). Temporal properties of the visual detectability of moving spatial white noise. *Exp Brain Res* **45**: 179–188.

Wichmann, F. A., & Hill, N. J. (2001a). The psychometric function I: fitting, sampling and goodness-of-fit. *Percept Psychophys* **63**(8): 1293–1313.

Wichmann, F. A., & Hill, N. J. (2001b). The psychometric function II: bootstrap based confidence intervals and sampling. *Percept Psychophys* **63**(8): 1314–1329.

Zeki, S. (2003). The disunity of consciousness. *Trends Cogn Sci* **7**(5): 214–218.

Zeki, S., & Bartels, A. (1998a). The asynchrony of consciousness. *Proc R Soc Lond B* **265**: 1583–1585.

Zeki, S., & Bartels, A. (1998b). The autonomy of the visual systems and the modularity of conscious vision. *Philos Trans R Soc Lond B Biol Sci* **353**: 1911–1914.

Zeki, S., & Bartels, A. (1999). Toward a theory of visual consciousness. *Conscious Cogn* **8**: 225–259.

Zeki, S., & Moutoussis, K. (1997). Temporal hierarchy of the visual perceptive systems in the mondrian world. *Proc R Soc Lond B Biol Sci* **264**(1387): 1415–1419.

Part IV

Spatial phenomena: forward shift effects

19

The Fröhlich effect: past and present

DIRK KERZEL

Summary

When observers are asked to localize the initial position of a moving target, they often indicate a position displaced in the direction of motion relative to the true onset position. In this review, the debate between Fröhlich, who discovered this phenomenon, and his contemporaries in the 1920s and 1930s is summarized. Striking misinterpretations of Fröhlich's findings and the anticipation of recent research on the flash-lag effect will be presented. In the second part, current accounts of the Fröhlich effect in terms of attention and metacontrast are evaluated. In the final section, reconciliation between research on the Fröhlich effect and recent reports of an error opposite the direction of motion (the onset repulsion effect) is offered.

19.1 Introduction

When asked to localize a moving target entering a window, observers often indicate a position not adjacent to the edge of the window but a position displaced in the direction of motion (see Fig. 19.1(a)). The gap between the edge of a window and the initial perception of the moving target was first discovered by the Norwegian astronomer O. Pihl in 1894, but Fröhlich (1923) was the first to study the effect systematically. Therefore, the illusion has been named the "Fröhlich effect." Fröhlich's explanation of the illusion in terms of "sensation time" was amply discussed in the 1930s (Fröhlich 1930, 1932; Rubin 1930; G. E. Müller 1931; Metzger 1932; Piéron 1935) but forgotten for the 60 years that followed. Research on the Fröhlich effect was revived at the end of the last century (Müsseler & Aschersleben 1998; Kirschfeld & Kammer 1999), and accounts of the phenomenon in terms of attention and metacontrast were forwarded. Yet more recently, an error opposite to Fröhlich's observation was reported (Thornton 2002), which is incompatible with all previous theories on the Fröhlich effect (Fig. 19.1(b)).

In the first section I will describe the methods, results, and theories of early research on the Fröhlich effect. It is surprising to see how much current work on the Fröhlich and flash-lag illusions was anticipated by past researchers and simply overlooked afterward. In the flash-lag illusion, a flash that is physically aligned with a moving object is perceived to lag behind (see Fig. 19.1(c)). I will also draw the reader's attention to the phenomenological aspects of stimulus localization described in detail in early research inspired by Gestalt

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

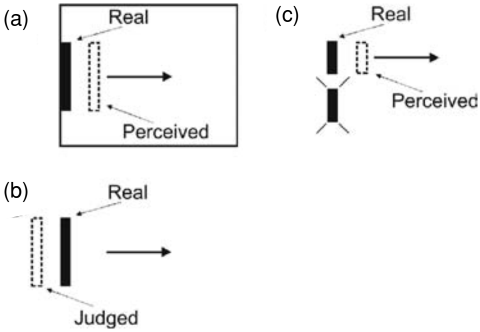


Fig. 19.1 The three illusions discussed in this chapter. In the Fröhlich illusion (a) the initial position of a bar entering a window is not perceived at the edge of the window, but some distance away from it. In the onset repulsion effect (b) the judged onset position is displaced opposite the direction of motion. In the flash-lag illusion (c) the position of a moving object that is physically aligned with a flashed object is seen ahead of the flash.

psychology. It shows that there may be more to localization then the report of a single position. In the second section, I will describe and evaluate current theories of the Fröhlich effect. In the final section, an attempt to reconcile the apparent contradiction between mislocalization of the initial position in the direction of motion (the Fröhlich effect) and recent reports of mislocalization opposite to the direction of motion (see Fig. 19.1(b)) will be presented.

19.2 Historical notes

In his seminal paper, Fröhlich (1923) reported not only one but several phenomena. In fact, Fröhlich's work did not focus on the phenomenon now considered to be the Fröhlich effect. Today, we consider the apparent displacement of the initial position of a moving target in the direction of motion the Fröhlich effect. Thereby, we refer to the displacement of the *trailing* edge of the target and denote that nothing is perceived between the physical onset position and the trailing edge, although the corresponding retinal positions were stimulated by the target. In contrast, Fröhlich was interested in the perceived position of the *leading* edge of the moving bar (see Fig. 19.2). He observed that the leading edge of the bar was not perceived right next to the border of the window and successively uncovered but appeared suddenly at a position displaced in the direction of motion relative to the edge of the frame. He considered the distance between the leading edge and the border of the frame an expression of the sensation time ("Empfindungszeit"), that is, the time between the impact of light and the corresponding visual sensation (Fröhlich 1923, 70–73). The position of the leading edge, x, divided by the velocity of the bar, equals the sensation time t, where $t =$ x/v. In Fröhlich's measurements, sensation time was found to be on the order of 100 msec with faint stimuli reducing to 50 msec with bright stimuli (see Fig. 19.2(c)).

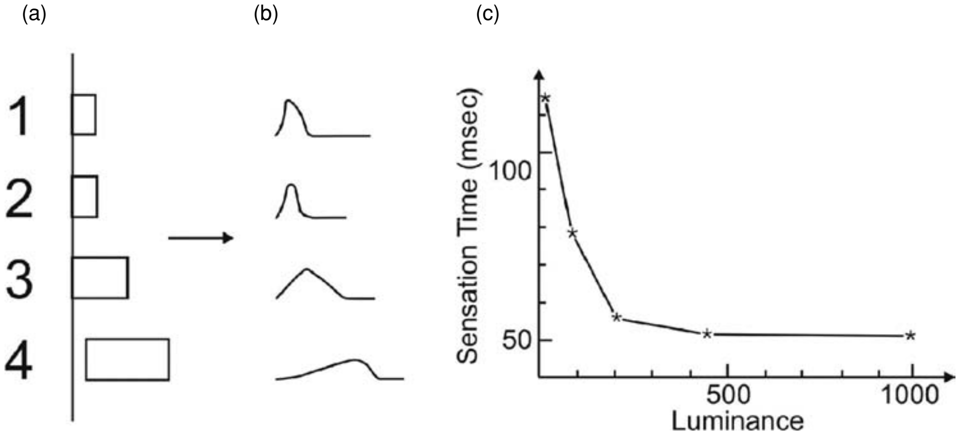


Fig. 19.2 Results of Fröhlich's (1923) experiments. (a) Appearance of a bar entering a window on the left and moving to the right according to Fröhlich (1923; adapted from p. 67). The bar appeared larger than it actually was and appeared in its enlarged width at once. The luminance of the bar decreases from (a)1 to (a)4. Only (a)4 corresponds to our current conception of the Fröhlich effect. (b) shows the perceived luminance distribution within the bars (adapted from p. 66). For the dim bar ((a)4), the luminance decreased toward its trailing edge and was highest at its leading edge. (c) Fröhlich thought that the displacement of the leading edge of the bar divided by its velocity indicated the sensation time. He measured sensation time as a function of luminance and noted that it was shorter for a bright bar (as in (a)1) than for a dim bar (as in (a)4). Luminance is given in proportional (but unknown) units (adapted from p. 74).

Fröhlich also noted that the perceived width of the bar was larger than that of the physical stimulus, which he interpreted as the perceived width capturing the duration of the primary sensation of the stimulus. On this view, the trajectory positions covered during this duration would be sensed at the same time. Because the perceived width of the bar decreased with increasing luminance, he concluded that luminance was negatively related to the duration of the primary sensation. From today's perspective, Fröhlich's ideas about the duration of the primary sensation seem untenable and have been supplanted with variable degrees of motion smear (Burr 1980) or visible persistence (Coltheart 1980).

Further, Fröhlich investigated the luminance distribution within the bar. He found that the bar looked brighter at its leading than at its trailing edge (see Fig. 19.2(b)). This effect was particularly pronounced with dim stimuli that appeared wider than bright stimuli. With dim stimuli, he also observed that the initial portion of the trajectory was not only darker than the leading edge but disappeared altogether. Fröhlich thought that the suppression of the initial part of the trajectory was due to the contrast arising between the leading edge of the bar and the previously covered positions. Thus, the phenomenon we consider to be the Fröhlich effect is only a limiting case of the larger class of phenomena observed by Fröhlich. The phenomenological observation of the luminance distribution within the target or the luminance distribution along the smeared out initial appearance of the target have been ignored in more recent research. This may be an error, because if observers are forced to report only a single point ("the onset") from a percept that shows graded levels of visibility or contrast, their judgments may be a matter of criterion. Observers have to decide what they should report and what is too weak to be worth reporting (Geer & Schmidt 2006). However, even if some of Fröhlich's phenomenological observations may be correct, his account of the illusion in terms of the "sensation time" is implausible, as contributions by his contemporaries show.

In a critique of Fröhlich's work, Rubin (1930) noted that reducing the size of the window and thereby shortening the visible trajectory of the moving bar reduced the Fröhlich effect (replicated in Müsseler & Neumann 1992). He concluded that a necessary condition for the occurrence of the Fröhlich effect was that the target continued to move after it entered the window. Fröhlich's account could only explain effects of trajectory length by assuming that the time at which a stimulus is sensed depends on processes occurring after the stimulus has been sensed. This is hard to maintain. As an alternative to Fröhlich's calculation of the sensation time, Rubin suggested that the distance between the position where the target appeared and the minimal trajectory length that resulted in a reduction of the Fröhlich illusion would be a better estimate of the sensation time. This distance indicated when perceptual processes influencing the appearance of the moving bar ended. Further, he noted that the magnitude of the Fröhlich effect was not determined by the absolute luminance as suggested by Fröhlich, but rather by the contrast between the moving element and the background.

To corroborate the hypothesis of continuing motion producing the Fröhlich effect, Rubin (1930) compared the perception of a stationary flash to the perception of the onset of a moving bar. To this end, he placed a narrow slit exactly above the edge of the elongated window where the target entered (see Fig. 19.3(b)). When the target line entered the narrow slit and the elongated window at the same time, the line in the window appeared displaced in the direction of motion relative to the slit. This, of course, is an early version of the flash-initiated cycle (Khurana & Nijhawan 1995) that was rediscovered in the debate on the flash-lag effect. Rubin also asked which of the two lines (the flashed or the continuously visible) was perceived first. Similar to more recent replications of this temporal order judgment (Nijhawan et al. 2004), he reported a lack of convergence between his own perceptions and those of his assistant. He concluded that the flash and the moving object appeared at about the same time).[1] Ironically, Fröhlich (1923) had run exactly the same experiment in his earlier publication but considered the displacement of the moving object relative to the flash and the perceived simultaneity of the two objects as support for his account.

Rubin (1930) further wondered whether the perception of the initial portion of the target was suppressed (as suggested by Fröhlich) or whether the sensations corresponding to the initial portion of the trajectory were displaced in the direction of motion. To test these conflicting possibilities, he presented a target that moved initially behind a red transparency

---

[1] In contrast to the perceived simultaneity of a flash and the onset of a moving stimulus, the temporal onset of a line moving at high speed precedes the onset of a stationary stimulus that stays on the screen (Kreegipuu & Allik 2003).

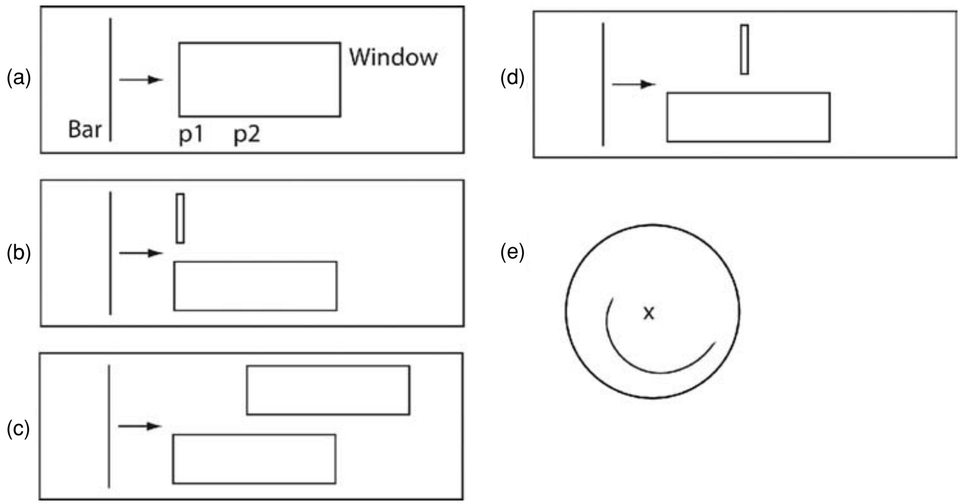


Fig. 19.3 Overview of experiments run in the 1920s and 1930s on the Fröhlich effect. (a) In Fröhlich's (1923) experiments, a bar moving behind a screen entered a window at position p1. The perceived initial position was read off a ruler below the window. Judgments of the leading edge of the bar were displaced in the direction of motion (position p2) (adapted from p. 101 in Rubin 1930). (b) Rubin (1930) and Fröhlich (1923) presented a second narrow window above the elongated window. They observed that the moving bar appeared ahead of the briefly flashed slit (adapted from p. 104 in Rubin 1930). Rubin asked observers to adjust a point below the window to the perceived initial position. (c) Metzger presented a bar moving through two elongated windows offset by a certain distance. The bar entering the window later (top window) appeared to lag behind the bar that was already visible (bottom window, adapted from p. 189 in Metzger 1932). (d) Metzger presented a bar moving through an elongated window and midway along the trajectory, the bar passed through a narrow slit above the elongated window. The flashed bar appeared to lag behind the moving bar below (adapted from p. 190 in Metzger 1932). Unfortunately, Metzger did not quantify his observations but only noted the direction of the displacement of the moving bar. About 90% of the observations that he collected from nine observers were consistent with the flash-lag effect. (e) Piéron presented part of a spiral on a rotating disk. If the disk were rotating clockwise, the spiral appeared further from the center than when it was rotating counterclockwise (adapted from Piéron 1935, p. 24).

and then behind a green transparency. If the initial stimulus characteristics were carried over to the positions further in the direction of motion, one would expect to see a red target on an otherwise green background. Rubin reported that observers' judgments were variable, but at least one subject reported seeing a red, stationary stimulus at the right place and then a green, moving object. This result suggests that the initial positions in the regular Fröhlich effect were suppressed rather than displaced into the direction of motion. In contrast to Rubin's suggestion, Cai (2003) reported that a red flash at the onset of motion was shifted in the direction of motion.

Metzger (1932) agreed with Rubin on the implausibility of Fröhlich's theory and suggested yet another one. His considerations focused on three types of appearance of the moving bar that showed overlap with the results of Fröhlich but were not quite the same.

The first type of appearance was observed at slow velocities. Metzger noted that even if the entrance point of the slit into the window were correctly perceived, the velocity of the bar seemed to change. It appeared slow at first and then accelerated after a while; a phenomenon later rediscovered by Runeson (1974). Second, Metzger observed (in agreement with Fröhlich) that the bar, at low contrasts, appeared suddenly in its entire length between the slit and the background. The perceived width of the bar exceeded its physical width when the bar first appeared, but as the bar started to move, the perceived width decreased. The third appearance type was a bar that appeared at a position offset from the edge of the screen, stood still for a moment, and then continued to move at a constant velocity. This type of appearance was more likely with high velocities and strong contrasts.

To explain these phenomena, Metzger (1932) suggested that sensation time was longest at the start of the motion and decreased as the motion progressed to a point where it became constant. He thought that the postulation of differential sensation times across the trajectory was almost trivial because "... every new process needs some time to 'shake down' and to 'push away' the process taking place at the same place..." (p. 185, translation by the author). According to Metzger, the assumption of differential sensation time could explain the three phenomena under investigation. For the sake of clarity, I will refer to the initial sensation time at the starting position p1 of the moving object as t1, and the final, constant sensation time at a position p2 further along the trajectory as t2 (see Fig. 19.3(a)). If the velocity of the target is slow such that the time needed to cover the distance between p1 and p2 is larger than the difference between t1 and t2, a moving target will be perceived with a velocity distortion (appearance type one). If the velocity of the target is intermediate, the time needed to move from p1 to p2 may equal the difference in sensation time (t1–t2) such that the complete trajectory between p1 and p2 reaches consciousness at the same time and a suddenly appearing, widened bar will be perceived (appearance type two). At high target velocities, the time needed to move from p1 to p2 may be far smaller than the difference in sensation time such that the position p2 may reach consciousness before position p1. In this case, both forward motion from p2 onward and backward motion from p2 to p1 would be perceived. Metzger noted that backward motion was mostly not perceived but that observers sometimes perceived a flicker. He argued that the conditions for the perception of the backward motion were unfavorable because it was dominated or masked by the much stronger forward motion.

To support his claims, Metzger (1932) conducted two experiments. First he placed two windows of unequal width above each other (Fig. 19.3(c)). A vertical bar moving behind the two windows was initially visible in only one of the windows. When the bar reached the edge of the other window, it became visible in the two windows. Although one and the same bar was viewed, the portion of the bar entering later appeared to lag behind and moved more slowly than the bar already visible. Second, he placed a small slit above the center of the window such that the moving bar would illuminate the vertical slit and the window at the same time and at the same horizontal position (Fig. 19.3(d)). The briefly illuminated slit appeared to lag behind the moving object. Metzger suggested that both phenomena were due to the longer sensation time at the beginning of a perceptual process. The initially longer sensation times explained why both the onset of a moving and a flashed stationary object appeared to lag behind a continuously visible moving object. Thus, Metzger not only discovered the flash-lag effect (that was rediscovered twice: MacKay 1958; Nijhawan 1994), but he also proposed differential latency (latency being a term comprising "sensation time") as an account of the phenomenon (for recent renditions see Purushothaman et al. 1998; Whitney & Murakami 1998). Although the differential latency account may justify the occurrence of the Fröhlich phenomenon and the flash-lag effect, it has difficulty accounting for the importance of the length of the trajectory (Rubin's first experiment). Accounts based on metacontrast overcome this limitation.

It was Piéron (1935) who first proposed that metacontrast masking was responsible for the suppression of the initial portion of the trajectory. Metacontrast masking was initially investigated by Stigler (1910) and refers to the fact that the visibility of a briefly flashed stimulus is reduced when it is followed by another stimulus in its spatial–temporal vicinity. The optimal stimulus onset asynchrony (SOA) between target and mask depends on the stimulus and task parameters but ranges between 40–100 msec; both shorter and longer SOAs reduce the masking effect. Piéron reasoned that the initial positions of a bar entering a window were masked by later presentations of the stimulus. Therefore, the initial portion of the trajectory was not perceived. In support of this idea, he presented a line on a rotating disk. The line approached the center of the disk while following its circumference (see Fig. 19.3(e)). When rotated counterclockwise, the distance between the line and the edge of the disk was larger than when the disk was rotated clockwise. In contrast, the distance between center and line was shorter when the disk was rotated clockwise, thus revealing the initial portion of the trajectory to be masked by subsequent stimulation.

Some 20 years later, Alpern (1953) pointed out the incompleteness of Piéron's account because it does not make clear why masking previous target positions stops at some point. If every target position along the trajectory masked the previously presented target positions, only the final target position should be visible. However, most of the trajectory is visible and only a small part at the beginning is invisible. Similarly, such a simple metacontrast account has difficulty explaining why the Fröhlich effect decreases with shorter trajectories. If only immediate neighbors mask the previous position,[2] then the number of positions following the initial position should not matter.

19.3 Cumulative lateral inhibition

The effect of trajectory length may be explained by assuming that inhibition from adjacent stimulus positions accumulates across the trajectory and is therefore stronger with longer trajectories (Geer & Schmidt 2006). To confirm this idea, Geer and Schmidt asked their

---

[2] The temporal separation between target and mask has a nonlinear effect on the strength of masking. Masking is maximal with SOAs of about 40–100 msec and decreases with shorter or longer intervals. Furthermore, masking decreases with spatial separation. These two factors (overview in Breitmeyer & Ögmen 2006) suggest that masking from successive positions of a moving stimulus will depend on target speed. However, these intricate interactions were not taken into account in early work on the Fröhlich illusion.

subjects to rate the brightness of selected positions along the target's trajectory. They found that the perceived contrast of the trajectory was weakest at the beginning and increased gradually. The increase of the perceived luminance along the trajectory was steeper when the trajectory was short, that is, the target reached maximal perceived contrast faster. Furthermore, the gradual increase in contrast permits observers to make a decision as to which part of the trajectory they report. When they adopt a more conservative criterion (report of only the high-contrast part of the trajectory) the Fröhlich effect was found to be larger than with a more liberal criterion. These recent experiments emphasize the importance of phenomenal aspects in the study of localization performance. In a situation with high uncertainty, the criteria adopted by the observers to interpret their percepts are key to understanding the nature of localization.

Nevertheless, the cumulative lateral inhibition account has the same difficulty as the simple metacontrast account in explaining why only the first positions of a moving object are invisible. What is lacking in all manner of metacontrast accounts is a component that determines when the target becomes visible again. In some accounts, this role is assigned to visual focal attention either with (Kirschfeld & Kammer 1999) or without (Müsseler & Aschersleben 1998) reference to metacontrast masking.

19.4 Attention shifting

Müsseler and Aschersleben (1998) proposed that the Fröhlich effect was the result of the time it takes to move focal attention to the moving stimulus to consciously perceive it. In general, the onset of a moving stimulus in the periphery elicits a shift of visual focal attention to this position. Visual focal attention greatly improves the speed and accuracy of visual information processing (Posner 1980; H. J. Müller & Rabbitt 1989) and may even be necessary for a stimulus to reach conscious awareness (Simons & Rensink 2005). During the time it takes the spot of attention to travel to the onset position of the moving target, the target moves away from its physical onset position. In a similar vein, Baldo and Klein (1995) suggested that the flash-lag effect was due to the time it takes to shift attention from the moving object to the flashed object. In the Fröhlich effect, the first position that benefits from enhanced processing through visual attention is displaced in the direction of motion. The attention-shifting account claims that the positions presented before the attention shift is executed are not perceived. The faster the target moves, the further it will move from its onset before attention reaches it. This idea predicts that the Fröhlich effect increases with increments in target velocity. This prediction has been largely confirmed (e.g., Fröhlich 1932; Kirschfeld & Kammer 1999; Kerzel & Müsseler 2002; Müsseler et al. 2002).

Further, it is expected that cueing the onset position of the moving target should reduce the Fröhlich illusion. A cue that precedes the target onset attracts visual focal attention and effectively reduces the time elapsed before the moving target is within the focus of attention. This prediction too has been confirmed: the Fröhlich effect was reduced when a cue was presented $\sim$120 msec before target onset in the vicinity of the initial target position (Müsseler & Aschersleben 1998; Kerzel & Müsseler 2002) or when a stationary cue was presented for 2.5 sec at the onset position (Whitney & Cavanagh 2000). In contrast, a cue of 2.5 sec did not affect the size of the flash-lag effect, showing that the Fröhlich effect and the flash-lag effect are distinct phenomena (for conflicting views see Metzger 1932; Eagleman & Sejnowski 2000). Similar to cueing, stimuli that allow for the efficient allocation of attention reduce the Fröhlich effect. For instance, attention is more easily focused on a single rotating dot compared to a rotating line passing through the fixation point (corresponding to a double cue, Posner & Cohen 1984), and the Fröhlich effect is larger for the harder to focus on stimulus (Kerzel & Müsseler 2002).

Finally, the attention-shifting account explains why mislocalization of the moving stimulus (flash-lag effect) is smaller in the complete-cycle relative to the flash-initiated cycle (Müsseler et al. 2002). At the beginning of the movement, attention is far from the moving object and a large mislocalization results. As the motion progresses, attention catches up with the moving object and the mislocalization is reduced.

Thus, on the plus side, the attention-shifting account accommodates effects of velocity, visual cues, and trajectory position. However, the relation between the magnitude of the attention shift and that of the Fröhlich effect is unclear. According to one view, attention travels with a constant velocity such that the time to complete an attention shift increases with distance (e.g., Posner et al. 1980; Posner & Cohen 1984). According to the contradictory view, attention shift is time invariant (Remington & Pierce 1984; Eriksen & Murphy 1987). If the duration of the attention shift increased with distance, one would expect the Fröhlich effect to increase with distance of the initial target position from the current focus of attention (the fovea in most cases). This view receives confirmation from the observation that the Fröhlich effect is larger for motion away from the fovea (i.e., when the distance increases after motion onset) than for motion toward it (i.e., when it decreases) (Müsseler & Aschersleben 1998). However, effects of eccentricity and in particular larger Fröhlich effects for more eccentric locations were not observed (Müsseler & Aschersleben 1998; Kerzel & Müsseler 2002). It is untenable that the Fröhlich effect depends on both distance-dependent and distance-independent attention shifts.

Further, the attention-shifting account claims that the initial positions of a moving object will only be available when the attention shift is complete. However, in conditions in which the Fröhlich effect was on the order of 2–3 deg, a slight change in the contrast of the moving stimulus made only 0.5 deg after the onset of motion was detected with 70% accuracy (Müsseler & Aschersleben 1998). If the initial portion of the trajectory never reached consciousness, how could the detection performance be so good? Additionally, the attention-shifting account predicts that regardless of the stimulus properties, the initial portion of the trajectory will be invisible. However, a colored flash at motion onset is "dragged" into the direction of motion and therefore cannot be considered invisible (Cai 2003).

19.5 Interplay between attention and metacontrast

Another model of the Fröhlich effect overcomes some of the lapses of the attention-shifting account by positing an interaction between visual focal attention and metacontrast (Kirschfeld & Kammer 1999). As already pointed out in the discussion of Piéron's (1935) work, the metacontrast account has problems explaining why we see more than the final position of a moving target. Remember that each presentation of the target leads to the suppression of previous target positions. To counteract this suppression, Kirschfeld and Kammer assumed that the onset of target motion elicited a shift of focal attention to the target and that visual attention was responsible for the visibility of the target. Similar to the attention-shifting account, it was assumed that the shift of attention takes some time, and before visual focal attention reaches the target, metacontrast has already suppressed the initial portion of the trajectory.

To test this view, Kirschfeld and Kammer (1999) investigated the localization of a rotating rod that was continuously illuminated but additionally flashed with far higher energy when it first appeared. The resulting percept was of a flashed bar at the correct initial position and a blurred bar that was displaced in the direction of motion (the Fröhlich effect). The interpretation of this striking phenomenon was that the transient, flashed illumination of the initial orientation was strong enough to overcome metacontrast masking, whereas the initial portion of the continuously lit bar was suppressed until focal attention arrived at the bar. Further, it was concluded that the moving bar had a shorter latency than the flashed bar, because the continuously visible bar appeared ahead of the flashed bar even though both bars had been presented simultaneously. Again, this condition replicates the results of Fröhlich (1923) and Rubin (1930) and repeats the idea that the spatial displacement may be used to estimate sensation time. However, Kirschfeld and Kammer hold that it only indicates the relative processing time of moving and flashed objects, not the absolute sensation time.

The approach that combines metacontrast masking and attention has the advantage that it easily accommodates the same findings as the attention-shifting approach (effects of velocity and cueing) and additionally explains why the initial portion of the trajectory is, in certain conditions, not completely invisible. If the features of the target at the beginning do not match its features during the rest of the trajectory (as in Müsseler & Aschersleben's [1998] detection experiment), masking may be reduced and the initial positions become visible again.

In this account of the Fröhlich effect, attention and metacontrast interact to produce the phenomenon. In studies unrelated to the Fröhlich effect, it was observed that attention may actually determine metacontrast masking (Di Lollo et al. 2000; but see Francis & Hermens 2002). Thus, attention and metacontrast are closely intertwined mechanisms; however, one may still question their harmony. On the one hand, the necessity of attention to travel to the target position explains why the Fröhlich effect increases with increasing velocity of the target. On the other hand, metacontrast is known to decrease with increasing distance between target and mask (Alpern 1953). As the distance between successive target presentations increases with increasing target velocity, this characteristic of metacontrast would actually predict a smaller Fröhlich effect at higher velocities. So far, such an inverted effect of velocity (i.e., a decreasing Fröhlich effect with increasing velocity) has not been observed.

One final problem with the attention-shifting and attention-shifting plus metacontrast approaches is that peripheral cueing does not completely eliminate the Fröhlich effect

(Müsseler & Aschersleben 1998; Whitney & Cavanagh 2000; Kerzel & Müsseler 2002). When attention is fully allocated to a particular position, processing of the moving object should be enhanced right from the start, thereby canceling the metacontrast-induced suppression. To defend the attention-shifting account, one may argue that the Fröhlich effect with peripheral cues persists because the shifts of attention into the periphery are sometimes incomplete due to the natural coupling between fovea and focus of attention. A further point against the attention-shifting account is that there is no evidence that distracting attention by an invalid cue increases the size of the Fröhlich effect as it should (Müsseler & Aschersleben 1998). Thus, predictions derived from the involvement of attention are again difficult to substantiate. Part of this problem may be the versatile, top-down and bottom-up nature of attention affected by an enormous number of factors. Thus, it is not always clear how and whether attention was actually modulated by an experimental manipulation. Most studies have failed to check attentional deployment using independent measures such as reaction times. A notable exception is a study by Khurana et al. (2000) that measured both attentional deployment via reaction times and the flash-lag effect but failed to find any effects of attention on spatial mislocalization.

19.6 Fröhlich effect versus onset repulsion

The studies reviewed so far unanimously report a localization error in the direction of motion. However, recently the opposite error has also been reported (see Fig. 19.4). That is, the onset position of a moving target was mislocalized opposite the direction of linear motion; the onset repulsion effect (ORE, first reported by Thornton 2002). In the case of curved trajectories, the ORE is opposite the tangents to the circular trajectory (see Fig. 19.4(c), Actis-Grosso & Stucchi 2003). It is evident that explanations in terms of attention shift, metacontrast, or sensation time do not apply to this error because the target is localized at a position it never occupied and perceptual processes were never triggered for these positions. In terms of velocity, the effects depend on the range of velocities presented in an experimental session. Thus, increased velocity renders either a greater ORE (Kerzel 2002; Thornton 2002) or does not have any effect at all (Hubbard & Motes 2002; Kerzel 2002; Actis-Grosso & Stucchi 2003). When the target velocity was drawn from a relatively slow range of velocities (∼5 to ∼20 deg/sec), effects of velocity were absent or reversed compared to the Fröhlich effect. However, when the range of velocities was expanded from ∼5 to ∼40 deg/sec, increasing velocity shifted the judged position toward the direction of motion (Kerzel & Gegenfurtner 2004). This is a first indication that the ORE is susceptible to the context across trials in an experiment; that is, the judgments in a given trial can be influenced by what is presented in other trials in the same session. This is not the case for the Fröhlich phenomenon. The forward error and the increase of the error with increasing velocity persist regardless of velocity range (Kerzel 2002).

Furthermore, the ORE depends on motion type. It is largest with smooth, continuous motion and decreases with implied motion (Thornton 2002; Kerzel 2004). In a sequence of implied motion, successive target presentations were separated by large spatiotemporal

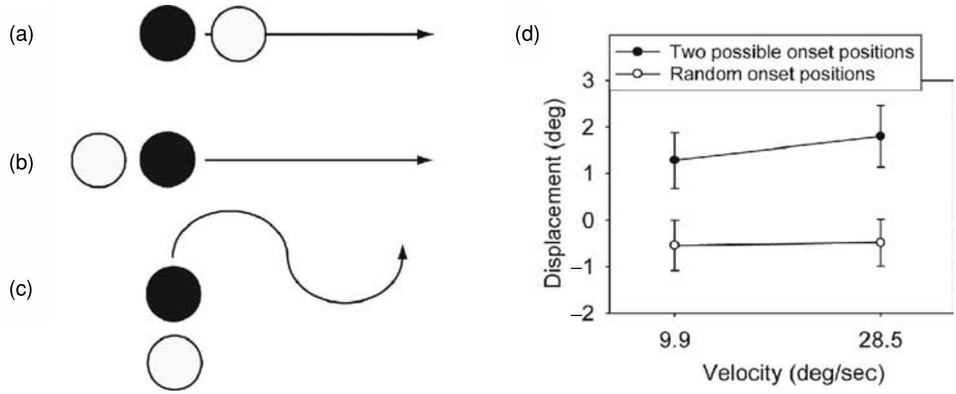


Fig. 19.4 The judged (open circle) and true onset position (filled circle) in the Fröhlich and onset repulsion effects. (a) In the Fröhlich effect, the onset position is displaced in the direction of motion. (b) In the onset repulsion effect (Thornton 2002), the onset position is mislocalized opposite the direction of motion. Most studies that reported an onset repulsion effect did not present visual references and used pointing movements. (c) In the case of circular trajectories, the initial position is displaced along the tangents of the trajectory (Actis-Grosso & Stucchi 2003). (d) By changing the uncertainty about where a target will appear, the Fröhlich effect may be turned into an onset repulsion effect. With two possible onset positions, judgments are displaced in the direction of motion and displacement increases with increasing target velocity. With random onsets, the forward error is eliminated (adapted from Müsseler & Kerzel 2004).

gaps such that each target position was more salient than with smooth motion. Because smooth target motion may elicit smooth pursuit eye movements, one may conjecture that the ORE is related to oculomotor control. This, however, was not the case as the ORE was not different in a condition with and without eye movements (Thornton 2002).

So what explains the difference between the ORE and the Fröhlich effects? The most likely reason has to be sought in the experimental procedure used to measure the error. In studies that have reported a Fröhlich effect, the onset position was judged relative to one or two environmental reference marks, such as the edge of a window (Fröhlich 1923; Piéron 1935; Kirschfeld & Kammer 1999), another moving target (Whitney & Cavanagh 2002), or two positions at a fixed eccentricity (Müsseler & Aschersleben 1998). In studies that have reported the ORE, such a fixed reference mark was missing. This was in particular the case in studies that used some form of pointing response (Hubbard & Motes 2002; Kerzel 2002; Thornton 2002; Kerzel & Gegenfurtner 2004; Müsseler & Kerzel 2004). When the target appears randomly in a relatively large area, and observers have to point to the onset position, the localization task effectively turns into an egocentric localization task. That is, observers have to localize the target with respect to their own body. In a direct comparison of pointing and relative judgments, Kerzel (2002) found that judgments of the onset position relative to a probe stimulus were displaced forward (Fröhlich effect), but mouse-pointing responses to the same stimuli were displaced backward (ORE). Thus, one may ask which attribute that distinguishes motor pointing and relative judgments accounts for the discrepant results.

A first hypothesis may be that the result of a perceptual comparison between moving target and probe stimulus is immediately available, whereas pointing movements are delayed and require memory of the initial position of the target after the trajectory has been viewed. In other words, the Fröhlich effect may be a perceptual effect while the ORE is based on memory. However, the Fröhlich effect has been observed with a probe stimulus that either appeared some time before or after target onset (Kerzel 2002). Thus, the Fröhlich effect is not only observed with immediate perceptual comparisons but persists in memory. Similarly, pointing movements render an ORE irrespective of whether responses are immediate or delayed (Kerzel & Gegenfurtner 2004). Thus, it is not the temporal aspect that is critical to the difference between relative judgments and motor pointing.

A more viable hypothesis considered by several authors (Kerzel 2002; Thornton 2002; Actis-Grosso & Stucchi 2003; Kerzel & Gegenfurtner 2004; Müsseler & Kerzel 2004) is that the uncertainty about the initial position causes observers to overcompensate for a potential error. When confronted with the task of localizing the initial position of a moving target, the most obvious error is to point to a position that is further along the trajectory. To avoid this, observers (perhaps unconsciously) compensate too much. To test this hypothesis, Müsseler and Kerzel (2004) investigated the localization of two positions at $\sim$7 deg of eccentricity to the left and right of fixation in two different trial contexts. In the random trial context, the target appeared mostly at a random position in a large area of the screen. Only in $\sim$17% of the trials did the target appear at the $\sim$7 deg positions. In the constant trial context, the target always appeared in one of the two eccentric positions to the left and right of fixation (similar to Müsseler & Aschersleben 1998). In both conditions, observers localized the onset position by using a mouse cursor. The results showed an ORE in the random context condition and the Fröhlich effect in the constant trial condition (see Fig. 19.4(d)). Thus, the high uncertainty about where a target will occur induces an error opposite the direction of motion with pointing tasks.

Another way to manipulate uncertainty in the presence of a reference object is to vary the distance between the onset position and the reference. Hubbard and Motes (2005) found a Fröhlich effect when the initial position of the target was adjacent to a large surrounding frame (similar to Fröhlich's window) and no or backward displacement when the initial position was far from the frame. Thus, it may be the availability of salient reference marks[3] for localization of the initial position that determines whether an error in or opposite the direction of motion will occur. Hubbard and Motes suggested that the reference frame provided the observer with a limit in their attempt to retrospectively reconstruct the trajectory. The frame (or occluding plane) offered a ready explanation why the target was not visible before its appearance. Without such a delimiting stimulus, observers may attempt to retrospectively extrapolate a possible prior trajectory of the target that appeared all of a sudden. This is particularly true when the onset is unpredictable and not salient (smooth motion as opposed to implied motion).

---

[3] Note that a structured background (Thornton 2002) or the presence of a ruler (Actis-Grosso & Stucchi 2003) that can be used to read the position of the target would not qualify as salient visual references because a background or a ruler does not provide a unique point of comparison.

Thus, the two localization errors that have been presented in this review are by no means contradictory. In conditions of high uncertainty, for instance during egocentric motor localization and in the absence of visual references, an error opposite the direction of motion occurs. This error is more or less constant across target velocities and highly susceptible to effects of across-trial context, which shows that it is related to observers' strategies (of error avoidance). This backward error may combine with the Fröhlich effect when an appropriate range of velocities is selected. Overall, an error opposite the direction of motion results, but increasing the target velocity shifts the judged initial position in the direction of motion. Thus, at some high velocity, the typical Fröhlich effect is replicated (Kerzel & Gegenfurtner 2004).

In sum, Fröhlich's observation that the initial portion of a moving target was invisible is currently explained by attentional latencies or the interplay between attention and metacontrast. The present review favors the latter explanation comprised of an inhibitory component (metacontrast) that explains why the initial portion is invisible, and a facilitatory component (attention) that explains why the trajectory becomes visible again. Nonetheless, evidence for the contribution of attention is mixed and requires further clarification. Further, the mislocalization of the onset opposite the direction of motion is related to uncertainty about target appearance resulting in observers overcompensating the distance traversed by the target after its appearance. The two errors are not contradictory, but rather complementary; the forward error reflecting perceptual while the backward error reflecting cognitive processes.

Future studies are needed to disentangle cognitive and perceptual components in the mislocalization of the initial position of a moving target. One route is to manipulate participants' strategies via feedback. It seems plausible that the magnitude of the ORE will change as a function of feedback, whereas the Fröhlich effect will not. Investigations that quantify the amount of metacontrast masking by successive target presentations along an object's trajectory are missing. As laid out above, a number of researchers have speculated about the involvement of metacontrast masking in the Fröhlich illusion, but there are no data relating masking functions in "static" target-mask displays to the Fröhlich effect. These data may also clarify which aspects of the pattern of mislocalization cannot be accounted for by low-level perceptual processing, but have to be attributed to higher-level functions such as attention.

Acknowledgment

D. Kerzel was supported by the Swiss National Foundation 10011–107768/1.

References

Actis-Grosso, R., & Stucchi, N. (2003). Shifting the start: backward mislocation of the initial position of a motion. *J Exp Psychol Hum Percept Perform* **29**(3): 675–691.

Alpern, M. (1953). Metacontrast. *J Opt Soc Am* **43**(8): 648–657.

Baldo, M. V., & Klein, S. A. (1995). Extrapolation or attention shift? *Nature* **378**(6557): 565–566.

Breitmeyer, B., & Öğmen, H. (2006). *Visual Masking: Time Slices Through Conscious and Unconscious Vision*. Oxford, UK: Oxford University Press.

Burr, D. (1980). Motion smear. *Nature* **284**(5752): 164–165.

Cai, R. (2003). The Fröhlich effect is not due to a failure to perceive the beginning portion of motion trajectory. *J Vis* **3**(9): 485.

Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics* **27**(3): 183–228.

Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: the psychophysics of reentrant visual processes. *J Exp Psychol Gen* **129**(4): 481–507.

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science* **287**(5460): 2036–2038.

Eriksen, C. W., & Murphy, T. D. (1987). Movement of attentional focus across the visual field: a critical look at the evidence. *Perception & Psychophysics* **42**(3): 299–305.

Francis, G., & Hermens, F. (2002). Comment on "Competition for consciousness among visual events: the psychophysics of reentrant visual processes" (Di Lollo, Enns, & Rensink, 2000). *J Exp Psychol Gen* **131**(4): 590–593; discussion 594–596.

Fröhlich, F. W. (1923). Über die Messung der Empfindungszeit. [On the measurement of sensation time]. *Zeitschrift für Sinnesphysiologie* **54**: 58–78.

Fröhlich, F. W. (1930). Über die Messung der Empfindungszeit. Eine Erwiderung auf experimenteller Grundlage [On the measurement of sensation time. An answer on an experimental basis]. *Psychologische Forschung* **13**: 285–288.

Fröhlich, F. W. (1932). Bemerkungen zu G.E. Müllers Kritik der Empfindungszeitmessung [Remarks on the criticism of G. E. Müller of the measurement of sensation time]. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane* **62**: 246–249.

Geer, M., & Schmidt, W. C. (2006). Perception of initial moving target signals: support for a cumulative lateral inhibition theory. *J Exp Psychol Hum Percept Perform* **32**(5): 1185–1196.

Hubbard, T. L., & Motes, M. A. (2002). Does representational momentum reflect a distortion of the length or the endpoint of a trajectory? *Cognition* **82**(3): B89–99.

Hubbard, T. L., & Motes, M. A. (2005). An effect of context on whether memory for initial position exhibits a Fröhlich Effect or an Onset Repulsion Effect. *Q J Exp Psychol* **58A**(6): 961–979.

Kerzel, D. (2002). Different localization of motion onset with pointing and relative judgements. *Exp Brain Res* **145**(3): 340–350.

Kerzel, D. (2004). Attentional load modulates mislocalization of moving stimuli, but does not eliminate the error. *Psychon Bull Rev* **11**(5): 848–853.

Kerzel, D., & Gegenfurtner, K. R. (2004). Spatial distortions and processing latencies in the onset repulsion and Fröhlich effects. *Vision Res* **44**(6): 577–590.

Kerzel, D., & Müsseler, J. (2002). Effects of stimulus material on the Fröhlich illusion. *Vision Res* **42**(2): 181–189.

Khurana, B., & Nijhawan, R. (1995). Extrapolation or attention shift. *Nature* **378**(6557): 566.

Khurana, B., Watanabe, K., & Nijhawan, R. (2000). The role of attention in motion extrapolation: are moving objects "corrected" or flashed objects attentionally delayed? *Perception* **29**(6): 675–692.

Kirschfeld, K., & Kammer, T. (1999). The Fröhlich effect: a consequence of the interaction of visual focal attention and metacontrast. *Vision Res* **39**: 3702–3709.

Kreegipuu, K., & Allik, J. (2003). Perceived onset time and position of a moving stimulus. *Vision Res* **43**(15): 1625–1635.

MacKay, D. M. (1958). Perceptual stability of a stroboscopically lit visual field containing self-luminous objects. *Nature* **181**(4607): 507–508.

Metzger, W. (1932). Versuch einer gemeinsamen Theorie der Phänomene Fröhlichs und Hazelhoffs und Kritik ihrer Verfahren zur Messung der Empfindungszeit [An attempt at a common theory of Fröhlich's and Hazelhoff's phenomena and a critique of their procedure for measuring sensation time]. *Psychologische Forschung* **16**: 176–200.

Müller, G. E. (1931). Erklärung der Erscheinungen eines mit konstanter Geschwindigkeit beewegten Lichtstreifens, insbesondere auch des Pihl-Fröhlichschen Phänomens [Explanation of the appearances of a lit bar moving at constant velocity, in particular the phenomenon of Pihl-Fröhlich]. *Zeitschrift für Sinnesphysiologie* **62**: 167–202.

Müller, H. J., & Rabbitt, P. M. (1989). Reflexive and voluntary orienting of visual attention: time course of activation and resistance to interruption. *J Exp Psychol Hum Percept Perform* **15**(2): 315–330.

Müsseler, J., & Aschersleben, G. (1998). Localizing the first position of a moving stimulus: the Fröhlich effect and an attention-shifting explanation. *Perception & Psychophysics* **60**(4): 683–695.

Müsseler, J., & Kerzel, D. (2004). The trial context determines adjusted localization of stimuli: reconciling the Fröhlich and Onset Repulsion Effects. *Vision Res* **44**(19): 2201–2206.

Müsseler, J., & Neumann, O. (1992). Apparent distance reduction with moving stimuli (Tandem Effect): evidence for an attention-shifting model. *Psychol Res* **54**(4): 246–266.

Müsseler, J., Stork, S., & Kerzel, D. (2002). Comparing mislocalizations with moving stimuli. The Fröhlich effect, the flash-lag effect and representational momentum. *Vis Cogn* **9**(1/2): 120–138.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature* **370**(6487): 256–257.

Nijhawan, R., Watanabe, K., Khurana, B., & Shimojo, S. (2004). Compensation of neural delays in visual-motor behaviour: no evidence for shorter afferent delays for visual motion. *Vis Cogn* **11**(2/3): 275–298.

Piéron, H. (1935). Le processus du métacontraste. *Journal de Psychologie Normale et Pathologique* **32**: 1–24.

Posner, M. I. (1980). Orienting of attention. *Q J Exp Psychol* **32**(1): 3–25.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. G. Bouwhuis (eds.), *Attention and Performance X*. Hillsdale, NJ: Erlbaum.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *J Exp Psychol* **109**(2): 160–174.

Purushothaman, G., Patel, S. S., Bedell, H. E., & Öğmen, H. (1998). Moving ahead through differential visual latency. *Nature* **396**(6710): 424.

Remington, R., & Pierce, L. (1984). Moving attention: evidence for time-invariant shifts of visual selective attention. *Perception & Psychophysics* **35**(4): 393–399.

Rubin, E. (1930). Kritisches und Experimentelles zur "Empfindungszeit" Fröhlichs [Critical and experimental remarks on Fröhlich's "Sensation Time]. *Psychologische Forschung* **13**(101–112).

Runeson, S. (1974). Constant velocity: not perceived as such. *Psychol Res* **37**(1): 3–23.

Simons, D. J., & Rensink, R. A. (2005). Change blindness: past, present, and future. *Trends Cogn Sci* **9**(1): 16–20.

Stigler, R. (1910). Chronophotische Studien über den Umgebungskontrast. *Pflügers Archiv für die gesamte Physiologie* **134**: 365–435.

Thornton, I. M. (2002). The onset repulsion effect. *Spat Vis* **15**(2): 219–244.

Whitney, D., & Cavanagh, P. (2000). The position of moving objects. *Science* **289**(5482): 1107a.

Whitney, D., & Cavanagh, P. (2002). Surrounding motion affects the perceived locations of moving stimuli. *Vis Cogn* **9**(1/2): 139–152.

Whitney, D., & Murakami, I. (1998). Latency difference, not spatial extrapolation. *Nat Neurosci* **1**(8): 656–657.

20

Approaches to representational momentum: theories and models

TIMOTHY L. HUBBARD

Summary

Memory for the final position of a target is usually displaced in the direction of target motion, a finding referred to as *representational momentum*. There are several different approaches to explaining representational momentum, and these approaches range from low-level perceptual mechanisms (e.g., oculomotor behavior) to high-level cognitive mechanisms (e.g., internalization of the effects of momentum). These approaches are overviewed, and a classification system involving internalization theories, belief-based theories, neointernalization theories, low-level theories, and network models is proposed. The extent to which each approach is consistent with the wide range of existent empirical data regarding representational momentum is noted, and possible directions of and considerations for a more unified theory of displacement are addressed.

Memory for the final position of a previously viewed target is often displaced in the direction of target motion. This forward displacement has been referred to as *representational momentum* (Freyd & Finke 1984) and is influenced by numerous variables (Hubbard 1995b, 2005). Although initial studies of representational momentum appeared consistent with the hypothesis that observers internalize or incorporate the principle of momentum into the representation of the target, subsequent studies reported displacement inconsistent with such a literal internalization or incorporation of momentum. For example, variables other than implied momentum such as conceptual knowledge about target identity (Reed & Vinson 1996), expectations regarding future target motion (Verfaillie & d'Ydewalle 1991; Johnston & Jones 2006), attributions about the source of target motion (Hubbard & Ruppel 2002; Hubbard & Favretto 2003), and whether observers visually track the target (Kerzel 2000; Kerzel et al. 2001) influence displacement. Displacement also occurs in the direction opposite to motion (Brehaut & Tipper 1996), along the axis orthogonal to motion (Hubbard & Bharucha 1988), and for stationary objects (Freyd et al. 1988). Any comprehensive approach to representational momentum needs to address this range of findings.

This chapter examines varied approaches to representational momentum, and several theories and models are discussed. Section 20.1 focuses on the definition of representational momentum and provides a brief overview of stimulus presentation, response collection, and empirical findings in studies of representational momentum. It is not an exhaustive review,

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

but rather is intended to provide background information (see more detailed review in Hubbard 2005). Section 20.2 provides descriptions of theories and models of representational momentum and displacement, introduces an organizational scheme for categorizing these theories and models, and notes whether each of the theories and models is consistent with, inconsistent with, or does not address the findings in Section 20.1. Section 20.3 considers the possibility of a more complete or unified theory of displacement and suggests considerations for a future theory of representational momentum and displacement. Section 20.4 provides a brief summary.

20.1 A brief overview of representational momentum

In order to understand theories and models discussed in Section 20.2, it is useful to define representational momentum, and then briefly review the methodologies and empirical findings from studies on representational momentum and related types of displacement.

20.1.1 Defining representational momentum

Momentum of a physical object is equal to the product of that object's mass and velocity (i.e., momentum $=$ mass $*$ velocity), and so representational momentum of a given target would presumably reflect the mental representation of the mass and velocity of that target. However, many researchers have used "representational momentum" in a broader sense to refer to any displacement in the remembered position of a previously viewed object that is in the direction of motion (or even any displacement in remembered position more generally). Hubbard (1995b, 2005) urged that "representational momentum" be used in a narrower sense to refer only to the component of displacement that reflected implied momentum, but such a narrower usage has not been uniformly adopted (e.g., see Thornton & Hayes 2004). The term "representational momentum" has also been used to describe both the pattern of displacement and a hypothesized explanatory mechanism for displacement. Although which meaning is intended is usually clear from the context, this dual usage can at times lead to confusion.

20.1.2 Methodology

In the representational momentum literature, stimuli can be presented in any of several different formats, and responses can be collected with any of several different measures. Even so, almost all studies of representational momentum involve computer-driven generation or presentation of stimuli and computer-assisted collection of responses using keyboard presses, cursor positioning and mouse clicks, or touching a computer monitor.

20.1.2.1 Stimulus presentation

Freyd and Finke (1984) presented observers with a small set of discrete concentric rectangular stimuli (inducing stimuli) that implied clockwise or counterclockwise rotation (see

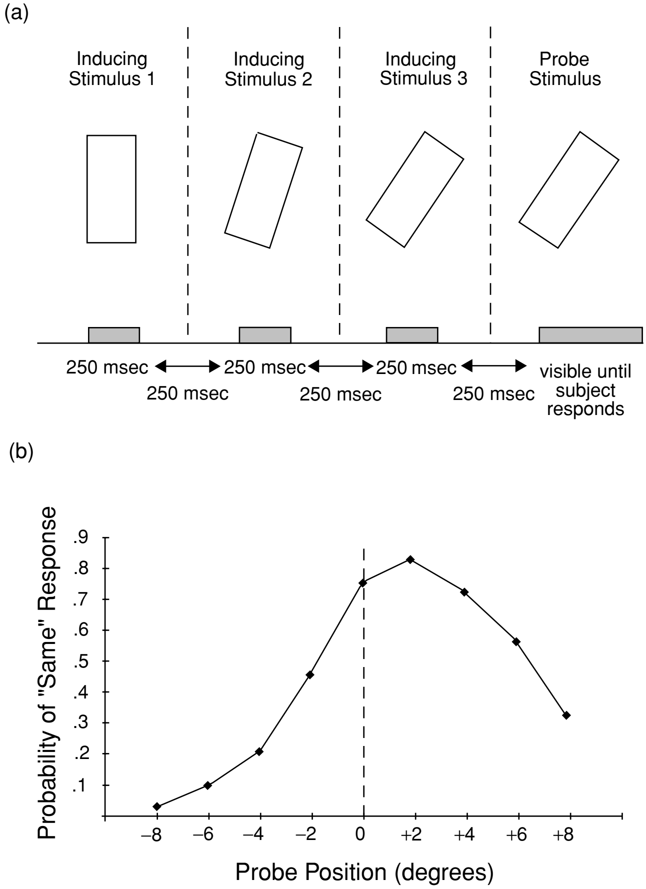


Fig. 20.1 The experimental methodology and results from Freyd and Finke (1984). (a) illustrates a typical trial in which three inducing stimuli and a probe are presented. (b) illustrates the probability of a *same* response as a function of probe orientation relative to the final inducing stimulus. The dashed line is the "true-same" orientation of the final inducing stimulus; negative probes were rotated backward from the orientation of the final inducing stimulus by the indicated number of degrees, and positive probes were rotated forward from the orientation of the final inducing stimulus by the indicated number of degrees. Representational momentum is indicated by the higher probability of a *same* response to positive probes.

top of Fig. 20.1). Hubbard and Bharucha (1988) presented observers with targets that exhibited continuous horizontal or vertical motion and vanished without warning.[1] Implied motion and continuous motion are the most common methods of stimulus presentation, but

---

[1] Motion depicted in computer-generated animation is never actually "continuous" or "smooth," as computer-generated motion involves discrete presentation of a target at one location followed by a redrawing of that target at a nearby location. However, if the separate presentations occur quickly enough, continuous and smooth motion is perceived. When researchers on representational momentum speak of "continuous" or "smooth" motion, they refer to displays in which differences between successive presentations are not perceivable, and so a target appears to exhibit continuous and smooth motion.

other methods have also been used. Freyd (1983) presented observers with single frozen-action photographs drawn from longer motion sequences (e.g., a person in mid-jump). In variations of this method, Freyd et al. (1988) presented drawings of stimuli that portrayed physical forces in equilibrium (e.g., a weight on a spring), and Hubbard and Courtney (2006) presented drawings of a dynamic figure (i.e., the T'ai-chi tu [yin-yang] symbol). In most studies, observers passively viewed targets, but Jordan and Knoblich (2004) gave observers partial control over direction and velocity of targets, and Jordan et al. (2002) gave observers partial control over when targets vanished.

20.1.2.2 Response collection

Freyd and Finke (1984) presented a probe stimulus after the final inducing stimulus vanished, and observers judged whether the probe was at the same position as the final inducing stimulus or at a different position. Over trials, the probe was slightly in front of, at the same position as, or slightly behind the actual position of the final inducing stimulus. The distribution of probe responses provides an estimate of displacement (see bottom of Fig. 20.1), and this typically involves calculation of either (a) the peak of a quadratic regression (e.g., Freyd & Johnson 1987), (b) the weighted mean (e.g., Munger, Solberg, Horrocks et al. 1999), or (c) the point of subjective equality (e.g., Kerzel 2003c). Probe judgment is the most common response method, but other methods have also been used. A more direct method of measuring displacement introduced by Hubbard and Bharucha (1988) involved observers using a computer mouse to place a cursor at the display coordinates where a target was judged to have vanished, and the difference between the judged vanishing point and the actual vanishing point provided a measure of displacement. More recently, researchers measured displacement using reaching responses in which observers touched the display at the coordinates at which a target was judged to have vanished (e.g., Kerzel & Gegenfurtner 2003; Motes et al. 2008).

20.1.3 Empirical findings

A wide range of variables influences representational momentum and related types of displacement, and Hubbard (2005) classified these variables as characteristics of the target, display, context, or observer.

20.1.3.1 Characteristics of the target

Forward displacement is usually greater with faster target velocities (Freyd & Finke 1985; Hubbard & Bharucha 1988), and accelerating targets exhibit greater forward displacement than do decelerating targets, even when final velocity is constant (Finke et al. 1986). When target motion occurs in the picture plane, forward displacement is greater for horizontal motion than for vertical motion (Hubbard & Bharucha 1988). Descending motion yields greater forward displacement than does ascending motion (Hubbard 1990; Munger & Owens 2004), but differences in forward displacement between leftward motion and rightward motion are not consistently observed (cf. Halpern & Kelly 1993; Hubbard & Bharucha

1988). Conceptual knowledge regarding target identity influences forward displacement (e.g., an upward moving stimulus labeled "rocket" exhibits greater forward displacement than does an otherwise identical stimulus labeled "cathedral," Reed & Vinson 1996; also Vinson & Reed 2002). The size or implied mass of the target does not influence forward displacement in the direction of motion (Cooper & Munger 1993), but larger targets exhibit greater downward displacement along the axis aligned with implied gravitational attraction regardless of the direction of target motion (Hubbard 1997). Forward displacement is observed for simultaneous multiple targets, even when each target moves in a different direction (Finke & Freyd 1985). Although most studies of forward displacement present visual stimuli, forward displacement has also been found with auditory (Freyd et al. 1990; Getzmann et al. 2004; Johnston & Jones 2006) and tactile (Brouwer et al. 2004) stimuli.

20.1.3.2 Characteristics of the display

If targets rotate in the picture plane, implied motion and continuous motion result in equal forward displacement (Munger & Owens 2004), whereas if targets translate in the picture plane, implied motion results in smaller (Faust 1990) or greater (Kerzel 2003c) forward displacement than does continuous motion. Effects of acceleration and deceleration of the target are greater with continuous motion than with implied motion (Poljansek 2002). Some investigators find forward displacement increases during the first few hundred milliseconds of the retention interval and then decreases with further increases in retention interval (e.g., Freyd & Johnson 1987), but only when target motion is highly predictable (e.g., Kerzel 2002a). Other investigators find an increase and asymptote in forward displacement with increases in retention interval (e.g., Kerzel 2000) or no effect of retention interval (e.g., Halpern & Kelly 1993) on forward displacement. Greater forward displacement is observed with motor responses such as reaching than with perceptual responses such as probe judgment (Kerzel 2003c; Kerzel & Gegenfurtner 2003).

20.1.3.3 Characteristics of the context

Forward displacement increases when a nearby or surrounding stimulus moves in the same direction as the target, and decreases when a nearby or surrounding stimulus moves in the direction opposite to the target (Hubbard 1993b; Whitney & Cavanagh 2002). Forward displacement increases if the target moves toward a landmark and decreases if the target moves away from a landmark (Hubbard & Ruppel 1999). Displacement also occurs along the axis orthogonal to target motion if that orthogonal axis is toward a landmark or other stimulus (Hubbard 1998b; Hubbard & Ruppel 1999) or is aligned with implied gravitational attraction (Hubbard 1990, 1997). If a nontarget stimulus is flashed near the end of target motion, forward displacement increases (Munger & Owens 2004), but if a nontarget stimulus is flashed during the retention interval between when the target vanished and a probe subsequently appeared, forward displacement decreases (Kerzel 2002b). Forward displacement is influenced by whether observers expect a target to change direction (Verfaillie & d'Ydewalle 1991; Hubbard 1994; Johnston & Jones 2006) and whether the final target location corresponds to a good or schematic ending (Hubbard 1993a; Kelly & Freyd 1987).

Forward displacement decreases if the target is initially stationary and subsequent motion of the target is attributed to impetus imparted from a moving stimulus that contacts the target and then launches that target into motion (Hubbard et al. 2001; Hubbard 2004).

20.1.3.4 Characteristics of the observer

Forward displacement increases when attention is divided between the target and another stimulus or task (Hayes & Freyd 2002; Joordens et al. 2004) and decreases when attention to the target is disrupted by presentation of a distractor during the retention interval (Kerzel 2003a). Forward displacement decreases but is not eliminated when the final location of the target is cued prior to when the target vanishes or during the retention interval (Hubbard et al. 2009) or when observers receive explicit instructions regarding representational momentum and are asked to compensate for its effects (Courtney & Hubbard 2008). Whether an observer tracks the target or fixates a stationary point away from the target influences forward displacement for continuously moving targets but not for implied motion targets (Kerzel 2000, 2003a; Kerzel et al. 2001). Forward displacement is influenced by activation of action plans (Jordan et al. 2002; Jordan & Knoblich 2004; Jordan & Hunsinger 2008), but is not influenced by whether observers receive feedback regarding the accuracy of their judgments of target position (Ruppel et al. 2009). Forward displacement is greater in younger children than in adults (Hubbard et al. 1999), but does not differ between older children and adults (Futterweit & Beilin 1994). Individuals with mental retardation exhibit smaller forward displacement (Jarrett et al. 2002), and individuals diagnosed with schizophrenia exhibit a trend for greater forward displacement (Conners et al. 1998) than do matched controls.

20.2 Theories and models of representational momentum

There are several theories and models of representational momentum (and of displacement more generally). Some address general properties of mental representation (e.g., Freyd 1987; Hubbard 2006a), whereas others address displacement for a specific stimulus type (e.g., continuous motion; Kerzel 2000), response measure (e.g., probe judgment; Bertamini 2002), or direction (e.g., forward; Kozhevnikov & Hegarty 2001). They can be grouped into five categories: (a) internalization theories, (b) belief-based theories, (c) neointernalization theories, (d) low-level theories, and (e) network models. The presentation here is necessarily brief and nonexhaustive, and general consistencies and inconsistencies of each theory and model with empirical findings noted in Section 20.1 are summarized in Table 20.1.

20.2.1 Internalization theories

Internalization theories suggest representational momentum results from properties of mental representation. The momentum metaphor is a specific theory regarding displacement in the direction of target motion, and spatiotemporal coherence is a general theory in which representational momentum is linked to changes in mental representation that result from

Table 20.1 *How well different theories of displacement account for the data*

	Theories of Displacement											
	INT		BB			N-INT		LL		NET		
	MM	SC	EE	IK	NI	SOI	AC	OB	MA	VA	EJ	BW
Target												
Velocity	+	+	+	+	+	+	+	+	+	+	+	+
Motion Direction	–	–	?	?	?	?	?	–	–	?	?	?
Orthogonal Direction	–	?	?	+	–	+	+	–	–	+	+	?
Identity	–	–	?	+	?	?	?	?	?	+	+	+
Multiple Targets	+	+	+	+	+	+	+	–	+	+	+	+
Mass/Size	–	?	?	?	–	+	+	?	?	+	?	?
Modality	+	+	+	+	?	+	+	–	?	+	+	+
Display												
Implied Motion	+	+	–	+	+	+	+	–	+	+	+	+
Continuous Motion	+	+	?	+	+	+	+	+	+	+	+	+
Frozen-action Photographs	+	+	–	+	+	+	+	–	–	+	+	+
Retention Interval	+	+	?	?	–	+	+	?	?	+	+	+
Cursor Positioning	+	+	?	+	+	+	+	+	–	+	+	+
Probe Judgment	+	+	–	+	+	+	+	+	+	+	+	+
Reaching	+	+	+	+	+	+	+	+	+	+	+	+
Context												
Surrounding Context	–	–	?	?	–	+	+	?	–	+	+	?
Landmarks	–	–	?	?	–	+	+	?	–	+	+	?
Future Motion	–	–	–	+	+	+	+	?	–	+	+	?
Schematic/ Good Ending	–	?	+	+	+	+	+	?	–	+	+	?
Attribution of Motion Source	–	–	?	+	+	+	+	?	–	+	?	?

Table 20.1 (*cont.*)

	Theories of Displacement											
	INT		BB			N-INT		LL		NET		
	MM	SC	EE	IK	NI	SOI	AC	OB	MA	VA	EJ	BW
Observer												
Allocation of Attention	–	?	?	?	?	?	?	+	+	+	+	+
Knowledge of RM/Feedback	?	?	–	?	?	?	?	?	–	?	?	?
Visual Tracking	–	?	?	–	–	?	?	+	+	?	?	?
Action Plans	–	?	?	?	?	?	+	+	–	+	+	?
Dissociation with Naive Physics	?	?	–	–	–	?	?	?	–	?	?	?
Age	?	?	?	?	?	?	?	?	?	?	?	?
Pathology	?	?	?	?	?	?	?	?	?	?	?	?

Note: A plus sign indicates data are consistent with or support a theory or model, a minus sign indicates data are inconsistent with or do not support a theory or model, and a question mark indicates data do not clearly address a theory or model. INT = internalization theories, BB = belief-based theories, N-INT = neointernalization theories, LL = low-level theories, NET = network models, MM = momentum metaphor, SC = spatiotemporal coherence, EE = explicit extrapolation, IK = implicit knowledge, NI = naive impetus, SOI = second-order isomorphism, AC = anticipatory consciousness, OB = oculomotor behavior, MA = motion aftereffect/perceptual adaptation, VA = vector addition, EJ = Erlhagen/Jancke model, BW = bow-wave model.

dynamic aspects of mental representation. Internalization theories were the earliest and most extensively discussed theories of displacement.

20.2.1.1 Momentum metaphor

The momentum metaphor suggested that the principle of momentum was incorporated into mental representations (Freyd & Finke 1984). As a consequence, mental representations exhibited a type of inertia. Just as a moving physical object cannot be immediately halted because of its momentum, so too a mental representation of that motion cannot be immediately halted because of an analogous momentum within the representational system (Finke & Freyd 1985; Finke et al. 1986). Finke et al. (1986) suggested the internalized form of momentum was relatively abstract, and so representational momentum could potentially accompany changes in stimuli that had no simple analogue to physical motion (e.g., changes in sound, size, color) if such changes could be extrapolated into the future; furthermore, such extrapolation could help observers (a) anticipate future positions of objects that move in a consistent manner, (b) regulate and control body movements, and (c) recognize objects moving to expected or familiar positions. Freyd et al. (1990) speculated representational momentum might have originated in the visual system as an internalization of momentum, but was "confiscated" by other neural systems to aid prediction of the future course of perceived events more generally. Although initially framed as an abstraction of change, the momentum metaphor has often been portrayed as a concrete internalization of physical momentum.

20.2.1.2 Spatiotemporal coherence

Freyd (1987, 1993) suggested representational momentum reflected spatiotemporal coherence between the represented and representing worlds (i.e., between the external physical world and the internal mental representation). Spatiotemporal coherence requires mental representation to be dynamic, and it is this dynamism that results in displacement. In order to be dynamic, a representation must intrinsically and necessarily include or incorporate time. An intrinsic inclusion or incorporation of time entails that the representation exhibits the same constraints as time; in other words, the representation of time must be directional and continuous. A necessary inclusion or incorporation of time requires that temporal aspects of represented information be an integral part of the representation and not a quality (or tag) distinct from a static and unchanging representation (i.e., the representation would need to systematically change over time). By displacing memory in the direction of target motion, and by exhibiting effects of retention interval, representational momentum reflects the directional and continuous aspects of dynamism, respectively. The necessary component of dynamism is reflected in findings that displacement of some sort occurs even if the magnitude and direction of displacement are influenced by observers' expectations or other knowledge (see also Finke & Freyd 1989).

Basing representational momentum on a broad conception of spatiotemporal coherence between the external physical world and the internal mental representation predicts that representational momentum should be found for any stimulus dimension that affords continuous change (Freyd 1992, 1993). This broad notion of spatiotemporal coherence was challenged by Brehaut and Tipper's (1996) finding that memory for the final luminance of inducing stimuli that increased or decreased in luminance was displaced backward toward an average of the inducing stimuli. Given that luminance is a perceptually continuous dimension, a broad spatiotemporal coherence notion predicts that forward displacement should have been exhibited. Brehaut and Tipper suggested representational momentum was limited to dimensions in which change entails motion, rather than being a general aspect of memory for any dimension that affords continuous change. Also, Hubbard (1999, 2006a) suggested spatiotemporal coherence reflected subjective aspects of physical principles rather than objective principles (as displacement reflects subjective experience of mass as weight, Hubbard 1997) and included effects of invariant physical principles in addition to momentum (e.g., gravity).

20.2.2 Belief-based theories

Belief-based theories suggest displacement results from beliefs regarding physical systems and objects in motion. A theory based on explicit extrapolation is a specific theory addressing displacement arising from implied motion, a theory based on implicit knowledge is a general theory, and a theory based on naive impetus is a specific theory focusing on aspects of forward displacement more consistent with notions of naive impetus than with the veridical understanding of momentum.

20.2.2.1 Explicit extrapolation

When discrete inducing stimuli are used to imply motion, it is possible that forward displacement results not from distortion in memory for the final inducing stimulus, but from observers predicting the position of a subsequent inducing stimulus. Finke and Freyd (1985) presented inducing stimuli consisting of patterns of dots in which each dot moved in a different direction. There was one forward probe (the configuration that would have occurred had the inducing stimuli continued) and one backward probe (the configuration identical to the previous [penultimate] inducing stimulus). A probe in which the configuration was the same as the final inducing stimulus was also presented. No difference in error rates or response times to forward probes and backward probes was exhibited, and there was no evidence of displacement. Finke and Freyd (1985) suggested this pattern demonstrated (a) observers in previous experiments were not predicting the position of a subsequent inducing stimuli, and (b) displacement in memory for final position did not result from masking of the final inducing stimulus by the probe or from observers confusing the final inducing stimulus with a previous inducing stimulus.

Finke and Shyi (1988) presented inducing stimuli similar to those in Finke and Freyd (1985). In a memory task, probes were clustered around the configuration of the final inducing stimulus, and forward displacement occurred. In an extrapolation task, probes were clustered around the configuration corresponding to the next configuration of inducing stimuli had the sequence of inducing stimuli continued, and marginally significant backward displacement occurred. Similar differences between displacement in a memory task and displacement in an extrapolation task (when three-dimensional renderings of three-armed figures exhibited apparent rotation), as well as increases in backward displacement in the extrapolation task with increases in target velocity, were reported by Munger and Minchew (2002). Furthermore, it was found that backward displacement in the extrapolation task increased with increases in target velocity. Given that the pattern of displacement in an extrapolation task differs from the pattern of displacement in a memory task, an account of displacement in memory for the final position of a target that is based on explicit prediction of the position of the next inducing stimulus can be rejected.

20.2.2.2 Implicit knowledge of physical principles

Hubbard (1998a) suggested representational momentum and related forms of displacement reflected implicit knowledge of physical principles. Because this knowledge is implicit, it is not available to explicit processes used in paper and pencil tests of physical knowledge (e.g., as in McCloskey & Kohl 1983), but might be available if observers judged veridicality of animated (i.e., dynamic) displays (e.g., as in Kaiser et al. 1985; Kaiser et al. 1992). This suggestion was meant to address an apparent contradiction between literature on naive physics (which suggested observers did not have veridical understanding of physical principles) and theories of displacement (some of which suggested observers did have veridical understanding of some aspects of physical principles). Furthermore, it was suggested that (a) if displacement was adaptive, then it should occur rapidly and without engaging attention or other cognitive resources (i.e., displacement should be automatic), and (b) one purpose of consciousness was to allow observers to learn and respond adaptively when mismatches between an automatically extrapolated position and subsequently sampled perceptual information occurred (i.e., when a target did not move as anticipated).

The idea that displacement involves implicit knowledge was bolstered by Freyd and Jones's (1994) observation that an observer's displacement pattern for targets ejected from a spiral tube did not correlate with that observer's performance on a paper and pencil test of explicit physical knowledge. Indeed, displacement patterns in Freyd and Jones were inconsistent with predictions based on veridical understanding of physical principles, as forward displacement was greater for targets that followed a physically incorrect spiral path after exiting a spiral tube than for targets that followed a physically correct straight path after exiting a spiral tube. Kozhevnikov and Hegarty (2001) reported that physics experts and physics novices exhibited greater forward displacement for smaller rising targets than for larger rising targets; however, physics experts correctly predicted larger objects would rise more rapidly than smaller objects, whereas physics novices incorrectly predicted smaller objects would rise more rapidly than larger objects. Kozhevnikov and Hegarty concluded that experts and novices had the same naive beliefs at the implicit level that influenced displacement, but that explicit physical knowledge could not penetrate the implicit level.

20.2.2.3 Naive impetus

Kozhevnikov and Hegarty (2001) suggested displacement attributed to representational momentum is more consistent with naive impetus theory than with Newtonian theory, and this is in accord with their findings regarding effects of target size on displacement of rising targets.[2] Additional evidence for the role of naive impetus in displacement was found in studies in which forward displacement of targets in launching effect displays (based on Michotte 1963) decreased relative to forward displacement of unlaunched control targets. According to naive impetus theory (McCloskey 1983), when a moving stimulus contacts a stationary target that results in the target moving, motion of the target is attributed to impetus imparted from the moving stimulus; an impetus that is believed to dissipate with subsequent target motion. If such a belief exists, then observers should expect the target to stop once the impetus dissipates below the threshold needed to maintain motion. Finke

---

[2] A theory of displacement based on naive impetus might seem just a special case of a more general theory of displacement based on implicit knowledge. For some observers, belief in naive impetus might specify the content of implicit knowledge; however, for other observers, belief in naive impetus might specify an explicit (but incorrect) physical principle those observers can clearly articulate (cf. McCloskey 1983).

et al. (1986) demonstrated forward displacement of a target decreased if observers expected the target to stop.

The idea that displacement reflects belief in naive impetus is consistent with the emphasis on the subjective consequences of physical principles on displacement in Hubbard (1999, 2006a). For example, when observers view an initially stationary physical object that begins moving immediately after being pushed, that pushed object subsequently slows and stops unless a compensating force (e.g., additional pushing) is applied. This reflects friction from the surface the object moves across or the medium the object moves through, but rather than mental representation incorporating the objective principle that an object in motion will continue in motion unless acted upon by an outside force and also incorporating the existence of friction as a separate outside force that acts upon a moving object, mental representation just incorporates the simpler impetus idea that an initially stationary object that began moving as a result of being pushed will slow and stop unless a compensating force is applied. The resulting idea of impetus could allow sufficiently accurate prediction of the behavior of most physical objects in most situations, and so observers could more easily model the behavior of targets by using a simpler (but incorrect) "impetus" notion than by using a more complex (and correct) "momentum plus friction" notion (Hubbard 2004).

20.2.3 Neointernalization theories

Neointernalization theories combine elements of internalization theories and belief-based theories, thus allowing both implied physical principles and the observer's expectations to influence displacement. A theory based on second-order isomorphism is a general theory in which displacement results in part from properties of the functional architecture of mental representation, and a theory based on anticipatory consciousness is a general theory in which displacement results from a remapping of perceptual space to reflect the observer's intended motor activity.

20.2.3.1 Second-order isomorphism

Hubbard (1999, 2006a) suggested displacement resulted in part from second-order isomorphism between invariant physical principles that operate on physical objects and mental representations of those objects. This notion can be illustrated by considering how representational momentum is consistent with Shepard's (1975, 1981; Shepard & Chipman 1970) notion of second-order isomorphism of objects and mental images of those objects. In physical rotation, a physical object at orientation A must pass through intermediate orientation B before reaching orientation C, and this reflects a constraint on physical transformations. In mental rotation, a mental representation of an object depicted at orientation A must pass through intermediate orientation B before reaching orientation C, and this reflects a constraint on mental transformations. The mental representation of the physical transformation is a functional analogue of the physical transformation, that is, mental rotation is second-order isomorphic to physical rotation. A physical object rotating from orientation A to orientation C would also exhibit momentum, and this reflects a constraint on physical transformation. A mental representation of a physical object depicted as rotating from orientation A to orientation C would also thus exhibit a functional analogue of momentum (i.e., representational momentum).[3]

A second-order isomorphism between subjective aspects of invariant physical principles that operate on physical objects and mental representations of those objects provides a default displacement consistent with a modified view of spatiotemporal coherence that emphasizes subjective aspects of invariant physical principles. In the absence of additional physical or cognitive context, this default displacement takes the form of a vector of activation that determines the direction and magnitude of displacement of the target (see Hubbard 1995b). The presence of additional physical or cognitive context provides additional vectors of activation that are added to the default displacement, and so the observed direction and magnitude of displacement for a given target reflects a combination of (a) the default displacement due to invariant physical principles, and (b) influences due to information provided by physical or cognitive context. By allowing physical or cognitive context to modulate the default displacement, such an approach resonates with the importance of context emphasized by Gestalt psychologists. Much as any given element of a display must be understood in terms of its relationship to other elements of that display, so too the displacement of any given target can only be understood in terms of that target's relationship to the context or event structure within which that target is embedded.

20.2.3.2 Anticipatory consciousness

Jordan's (1998) discussion of a possible anticipatory role of consciousness in perception suggests displacement reflects the intentions and anticipations of an observer. That is, displacement results from remapping the perceptual space of an observer to reflect that observer's intentions and anticipations. This remapping is a natural consequence of a common coding structure in which action planning and perception share neural resources, and the represented location of the target reflects the ongoing interaction between the observer's action plans and the actual stimulus location rather than reflecting just the actual stimulus location. Furthermore, Jordan (1998) suggests that dynamics of the environment have resulted in sensory-motor coordination that reflects those dynamics; more specifically, that dynamics of the environment have been transferred, via natural selection, into the algorithms of sensory-motor control. As a result, the location of the target and the location of the observer are contextualized by the target's "anticipated" location. Such internalization is consistent with the shaping of the functional architecture of mental representation suggested by second-order isomorphism (Hubbard 2006a) and with feedforward modeling in which perceptual encoding reflects consequences of potential actions (e.g., Desmurget & Grafton 2003).

---

[3] There are other intriguing connections between imagery and representational momentum. Kelly and Freyd (1987) speculated representational momentum reflects analogue representation similar to that suggested to underlie imagery. Munger, Solberg, and Horrocks (1999) reported observers who exhibited greater representational momentum, exhibited faster mental rotation, and suggested observers "filled in" between inducing stimuli with processes used in mental imagery. Senior, Barnes, and David (2001) reported participants with higher scores on the Vividness of Visual Imagery Questionnaire exhibit greater representational momentum. Hubbard (2006a) speculated mental imagery exhibits the same directionality and continuity that characterize the spatiotemporal coherence that Freyd (1987) hypothesized to give rise to representational momentum.

20.2.4 Low-level theories

Low-level theories suggest displacement results from properties of low-level vision such as oculomotor behavior, motion aftereffects, and perceptual adaptation. These narrow explanations do not appeal to knowledge or experience beyond that arising from perception of the current target, and suggest displacement does not involve or result from memory or other high-level cognitive processes. A theory based on oculomotor behavior originally appeared intended as a general theory (e.g., Kerzel 2000; Kerzel et al. 2001), but it only applies to targets exhibiting continuous motion (e.g., Kerzel 2003a,b). A theory based on motion aftereffects and perceptual adaptation is a specific theory applied to displacement measured by probe judgment.

20.2.4.1 Oculomotor behavior

Kerzel (2000; Kerzel et al. 2001) noted that pursuit eye movements overshoot the final position of a continuously moving target, and coupled with findings that memory for the position of a target is biased toward the fovea (e.g., Müsseler et al. 1999), it was suggested that forward displacement reflects movement and position of the eyes. Also, Kerzel (2000) noted visual persistence of a target (subjectively visible for 50–60 milliseconds after the target objectively vanished), coupled with the tendency for pursuit eye movements to overshoot the final position of a continuously moving target, suggests forward displacement reflects properties of the eyes and eye movements. Kerzel (2002c) presented probes during the brief interval in which visual persistence occurred, and the point of subjective equality in judgments of the alignment of the probe and the target was shifted forward in the direction of motion. It was also suggested that predictive eye movements accounted for changes in displacement accompanying expected changes in target direction (Kerzel 2002c). Furthermore, displacement previously attributed to representational friction (Hubbard 1995a; see Kerzel 2002c) or to representational centripetal force (Hubbard 1996; see Kerzel 2003b) is decreased or eliminated for continuous motion targets when observers cannot track the target.

Given the consequences of pursuit eye movements, foveal bias, and visual persistence, Kerzel (2000, 2002c; Kerzel et al. 2001) suggested forward displacement resulted from oculomotor behavior and was at least partly perceptual. In support of this, Kerzel (2002b) reported memory averaging of a target and nontarget stimulus occurred only if the nontarget stimulus was present when the target vanished or shortly thereafter. Indeed, Kerzel (2002c, p. 692) claimed "perceptual factors account for a large proportion of a mislocalization that was previously thought to result from processes operating in memory." Even so, an apparent challenge to an oculomotor behavior theory is that pursuit eye movements are not evoked by implied motion or frozen-action stimuli. In the absence of pursuit eye movements, oculomotor behavior cannot account for displacement with implied motion or frozen-action stimuli. However, Kerzel et al. (2001) argued fixation might still be shifted in the direction of motion even in the case of implied motion stimuli. Later, though, Kerzel (2003b; see also Kerzel 2005) suggested oculomotor behavior cannot be the only source of displacement, and that displacement with implied motion stimuli reflects high-level factors.

20.2.4.2 Motion aftereffects/perceptual adaptation

Bertamini (2002) suggested the greater likelihood of a *same* response to probes beyond the final position of a moving target reflected motion aftereffects and perceptual adaptation. If observers view motion in a specific direction, then a motion aftereffect will raise their threshold for detecting subsequent motion in that direction (Bonnet et al. 1984). Such observers might be less likely to perceive probes slightly beyond the actual final position than probes slightly behind the actual final position as different from the actual final position of the target. Similarly, if observers attend to a specific direction of motion, then they exhibit less sensitivity to subsequent events in that direction (Raymond et al. 1998). Thus, forward displacement results from asymmetrical change in sensitivity rather than from anticipation regarding the target. Bertamini's account stresses passive loss of sensitivity rather than active production of displacement, but it is not clear whether a passive loss of sensitivity is completely consistent with observations by Freyd and Finke (1984) and Hubbard and Bharucha (1988) that motion aftereffects are in the direction opposite to representational momentum and so could not actively produce representational momentum.

20.2.5 Network models

Although internalization, belief-based, and neointernalization theories specify possible cognitive mechanisms involved in displacement, those theories do not explicitly address implementation. Network models offer ways in which cognitive mechanisms of displacement might be implemented, and are typically general models that suggest displacement results from properties of network representations of the target and of the context within which the target is embedded (i.e., displacement results from patterns of connectivity and spreading activation within a network architecture). Given that other approaches discussed in Section 20.2 involve different levels of explanation than do network models, network models should be viewed as complementing rather than competing with other approaches.

20.2.5.1 Vector addition (weighted averaging)

One way to characterize how influences of implied physical principles and influences of physical or cognitive context contribute to displacement is to consider individual influences as separate vectors (Hubbard 1995b). The direction and magnitude of each vector reflects properties of the functional architecture of representation or information regarding the observer's knowledge or beliefs, and the displacement of a given target reflects a combination (e.g., summation or weighted average) of these vectors. Such vectors can be broadly construed as corresponding to magnitudes and directions of activation within a network architecture that preserve functional mapping between physical space and represented space. As a target moves through space, it traces a pattern of activation through the network. Nontarget context such as nearby surfaces and objects activates network locations corresponding to those surfaces and objects. Once a network location is activated, that activation spreads along excitatory or inhibitory pathways to neighboring locations. The remembered location of a target corresponds to the center of activation attributed to that target, and the difference between the center of activation and the network location corresponding to the veridical location of the target determines the direction and magnitude of displacement.

Hubbard (1995b) suggested how different influences on displacement might be modeled in a network architecture. Implied physical principles might be modeled by having activation channeled in specific directions (e.g., representational momentum as stronger excitatory activation along the anticipated path of motion, representational gravity as stronger flow of activation downward, and representational friction as inhibitory activation from a contacted region). Memory averaging might be modeled by residual activation from previous locations of the target, or by spreading activation from nearby nontarget stimuli, combining with activation from the final location of the target. This additional activation would shift the averaged center of activation representing the target toward those previous locations or nontarget stimuli. An observer's expectations could provide excitatory activation (priming) to network locations that corresponded to an anticipated position or provide inhibitory activation to network locations that did not correspond to an anticipated position. As noted earlier, the displacement of a given target would reflect some combination of all of these (and possibly other) influences.

20.2.5.2 Erlhagen-Jancke model

Erlhagen and Jancke (1999, 2004) developed a mathematical model generally consistent with speculations in Hubbard (1995b). One mechanism in the model involves interacting excitatory or inhibitory cell populations. Localized activity corresponds to the represented stimulus, and this activity moves through the network. Recurrent interactions within the network develop a wave pattern that sustains dynamic transformation (i.e., prolongs changes in patterns of activation) for a brief time after stimulus offset. After stimulus offset, excitatory activity continues to increase, reaches a maximum, and then decays back to a resting level. A second mechanism in the model involves cognitive factors (e.g., prior knowledge, action plans) that are modeled as additional dynamic inputs and influence the extent to which the cell population response overshoots final target position. If the threshold for gating recurrent interactions is low, there is larger extrapolation of past information into the future, and forward displacement of represented position occurs. If the threshold for gating recurrent interactions is high, as would happen if observers expect a target to reverse direction, then backward displacement in represented position occurs.

The Erlhagen-Jancke model addresses several findings in the displacement literature. First, the model addresses forward displacement with continuous motion and with implied motion. With continuous motion, continuous sampling produces a traveling wave of activity that produces displacement. However, with implied motion, the interstimulus intervals typically used (250 milliseconds) are too long to permit generation of such a traveling wave, and strictly bottom-up activity would not produce displacement. Erlhagen (2003) suggests that in such circumstances a bottom-up signal is continuously compared with an internal model that predicts future states of the moving stimulus, and this top-down influence bridges the interstimulus intervals between inducing stimuli, thus allowing a traveling wave capable of producing displacement to be generated. Second, the model addresses changes in displacement when observers expect a target to bounce off a barrier by incorporating hyperpolarizing units coding for position in the area of the barrier (thus raising the threshold), and so recurrent interactions are not sufficiently strong to allow forward displacement. Third, the model addresses effects of oculomotor action plans by adding a predictive signal that specifies the direction of an observer's gaze.

20.2.5.3 Bow-wave model

Müsseler et al. (2002) proposed the bow-wave model to account for representational momentum, the Fröhlich effect (forward displacement of the remembered initial location of a moving target; Müsseler & Aschersleben 1998), and the flash-lag effect (a briefly presented [flashed] stationary object aligned with a moving stimulus is perceived to lag behind the moving stimulus; Nijhawan 2002). The bow-wave model assumes a moving target produces activation within a network representing spatial location. Adjacent positions in the direction of movement are differentially primed by spreading activation from the target (cf. Hubbard 1995b). As the target approaches specific positions along its path, additional activation at those positions is accumulated, and a stimulus-driven "bow-wave" of activation occurs and spreads forward (cf. Erlhagen & Jancke 1999, 2004). A Fröhlich effect occurs because initial movement of the target skews activation forward; representational momentum occurs because activation requires time to decay, and during the course of decay, remaining activation continues to spread forward; and the flash-lag effect occurs because neural processing of a moving stimulus is faster than neural processing of a stationary (and flashed) target, and so the center of activation for the moving stimulus is displaced ahead of the center of activation for the flashed target.

20.2.6 Evaluating the theories and models

The theories and models discussed in Section 20.2 address different, and often nonoverlapping, types of data. Many specific theories and models arose from critiques of earlier studies or from criticisms of general theories and models, and might not have been intended as general accounts of displacement. Even so, if such specific theories or models are valid, they would presumably be consistent with other findings. Thus, it is useful to evaluate how specific theories and models, as well as how general theories and models, account for displacement over a wide range of experimental data. Such a summary is provided in Table 20.1, and whether each of the theories and models discussed in Section 20.2 is consistent with, inconsistent with, or doesn't address each of the experimental findings mentioned in Section 20.1 is noted. Effects of some variables (e.g., velocity) are consistent with many of the theories and models. However, effects of other variables (e.g., expectations regarding future motion, fixation away from a smooth motion target) are consistent with some theories and models and inconsistent with others. Effects of still other variables are not addressed by each of the theories and models, and it is not clear whether those effects are consistent or inconsistent. Neointernalization theories and network models seem consistent with more data than are other approaches, but there is still a significant amount of data that each theory or model does not address.

20.3 Toward a future theory of displacement

As shown in Section 20.2, there are a variety of theories and models regarding displacement and representational momentum, but no current theory or model adequately addresses all of the data on displacement and representational momentum, nor do investigators agree on what is the best approach. Some theories address very specific low-level mechanisms (e.g., pursuit eye movements and visual persistence), whereas other theories address more general high-level mechanisms (e.g., beliefs regarding a target or the operation of physical systems). What would be more useful is a more unified theory that addresses a broader range of data, and there are at least two different ways to approach the development of such a theory. One way is to integrate existing approaches into a larger, more inclusive theory, and this would be a more bottom-up approach (e.g., Kerzel 2005, 2006). A second way is based upon development of a computational theory of displacement, and this would be a more top-down approach (e.g., Hubbard 2005, 2006b).

20.3.1 Attempts at unification

Although some of the theories reviewed in Section 20.2 could be considered general theories intended to apply to a broad range of displacement data, none of those could be considered fully unified theories. By combining internalization and belief-based approaches, neointernalization theories provide an initial step toward a broader and more unified theory, but many aspects of the data nonetheless remain unaddressed (e.g., many characteristics of the observer). Further development of network models might produce a broader and more unified theory, and although such development could be useful for prediction and modeling, it is not clear that such models would be truly explanatory. There have recently been two suggestions regarding a more unified theory: Kerzel's three-factor approach and Hubbard's computational theory approach.

20.3.1.1 Kerzel's three-factor approach

A bottom-up approach to developing a more general theory of displacement might consider taking some of the separate theories discussed in Section 20.2, and then combining those to produce a larger and more comprehensive theory. Such an approach was taken by Kerzel (2005, 2006) in proposing a three-factor approach that brought together results from his work on effects of eye movements, response type, and motion type. This approach suggests the occurrence of forward displacement is determined by a combination of (a) the presence or absence of eye movements, (b) whether the response is verbal or motoric, and (c) whether target motion appears smooth and continuous or appears discrete and implied. Kerzel suggests pursuit eye movements contribute to forward displacement for targets exhibiting smooth motion, but not for targets exhibiting implied motion. Similarly, forward displacement can be produced by the motor system when a person reaches to indicate target position, but not when a person renders a verbal judgment regarding whether a subsequently presented probe is at the same position where the previously viewed target vanished. In the three-factor approach, the mechanisms that produce forward displacement exist primarily at the perceptual or motoric level rather than at the cognitive level, and there is no internalization or incorporation of physical principles.

Although the three-factor approach addresses displacement in a wider range of stimulus and response types than do any of the three factors taken individually, it does not address some key displacement data (e.g., downward displacement along the orthogonal axis of horizontally moving targets, forward displacement in frozen action photographs). More critically, some factors are treated as causal (rather than mediating or moderating) while other (often higher-level) factors might be more properly regarded as causal. For example, pursuit eye movements are treated as causal of forward displacement when observers view continuous motion (see Kerzel 2000), but it seems more parsimonious that causality should be assigned to high-level expectations regarding continuation of motion (and that drive such predictive or anticipatory eye movements) given that high-level expectations are already evoked to explain displacement when observers view implied motion. Similarly, differences in displacement related to expectations regarding changes in target behavior (e.g., direction of motion) are more parsimoniously attributed to high-level expectations than to multiple different low-level mechanisms (cf. Kerzel 2002; but see Hubbard 2006b). Also, by positing separate and unrelated mechanisms for the same displacement pattern in different types of stimuli, displacements resulting from different types of stimuli are viewed as separate phenomena, making it more difficult to determine appropriate generalizations or constraints.

20.3.1.2 Hubbard's computational theory approach

A top-down approach to developing a more general theory of displacement might focus on the function of displacement and on the information processing benefits of displacement. Hubbard (2005, 2006a) proposed that a theory of representational momentum should focus on a computational theory of displacement rather than on implementation per se. The computational theory of a process (or function more generally) addresses how that process helps an organism and what problem that process solves (Marr 1982). A computational theory of representational momentum would begin by considering how displacement might help an organism. Given that forward displacement resulting from implied motion or from continuous motion anticipates the behavior of a target, such displacement could be useful in spatial localization of stimuli in the environment, navigation, and survival. For example, a predator among shadows and occlusions in a jungle is visible only intermittently (similar to implied motion), but a predator in open grassland is visible continuously (similar to continuous motion). In both jungle or grassland settings, accurate anticipation of a predator's location would be adaptive, and forward displacement in the represented location of a predator in either setting would serve the purpose of aiding survival.

A computational theory addresses constraints on the information being computed, but does not address specific mechanisms regarding how the computation is implemented (e.g., see Dawson 1998). Indeed, given that different types of information are available in different settings, it would not be surprising if displacement in different settings were implemented in different ways. An approach based on computational theory views forward displacement resulting from implied motion or continuous motion as examples of the same phenomenon, and so seeks a deeper and broader level of explanation that would then guide a subsequent search for and interpretation of the specific representations, algorithms, and implementations of displacement. By focusing initially on the "big picture" rather than on the specific details of representation, algorithm, and implementation, such an approach is more likely to allow discovery of relevant generalizations and constraints.

20.3.2 Considerations for a future theory

There are several considerations for any future (and more general or unified) theory of displacement and representational momentum. Any future theory should address the full breadth of data and the multiple levels of processing and representation involved in displacement and representational momentum. Given that both low-level variables and high-level variables influence displacement, any future theory should address how information from one level can influence or be integrated with information from another level. Finally, any future theory should explicitly address the function of displacement, as such a consideration could help elaborate a subsequent computational theory of displacement.

20.3.2.1 Breadth of data

As noted in Section 20.2, many of the earlier theories of representational momentum were specific theories that addressed displacement resulting from a specific type of stimulus or response. Such specific theories tended to ignore the existence of displacement resulting from other types of stimuli or responses, and as a result, researchers sometimes appeared to not realize the limitations of those specific theories. Along these lines, consideration of displacement should not be limited to just forward displacement along the axis of motion, but should include other potentially related displacements (e.g., downward displacement attributable to implied gravity, displacement toward a nearby landmark).[4] Similarly, consideration of forward displacement should not be limited to effects of momentum per se, but

[4] It is unclear whether the types of displacement related to representational momentum might be related to other types of displacement such as Fröhlich effect, onset repulsion effect, or flash-lag effect. There have been preliminary attempts to account for these different types of displacement as well as representational momentum within a common mechanism (e.g., Müsseler et al. 2002) and to specify the relationship of representational momentum to these other displacements (e.g., Munger & Owens 2004; Hubbard 2006a), but more work remains to be done.

should address other variables that might also influence forward displacement. Examination of Table 20.1 suggests previous theories often addressed just one or two of the four types of characteristics (target, display, context, observer), but any future theory should address all four types of characteristics. In general, high-level theories (or theories involving high-level mechanisms) appear to have greater breadth than do low-level theories. Similarly, network models appear to have greater breadth, but this might just reflect their greater level of abstraction.

20.3.2.2 Multiple levels

Marr (1982) argued that a complete understanding of a given process would require understanding at each of the levels of computational theory, representation and algorithm, and implementation. Many of the theories in Section 20.2 focused on the levels of implementation or of representation and algorithm, and it is not clear whether a larger theory cobbled together from separate and specific theories of implementation or of representation and algorithm would sufficiently address issues at the computational theory level. Even if a Marr framework for displacement is rejected, any future theory of displacement must still address the issue of multiple levels of representation, as empirical findings regarding effects of both low-level variables and high-level variables on displacement suggest the existence of multiple levels of representation. The presence of low- and high-level variables is consistent with the existence of other hierarchical structures in perceptual (e.g., Treisman & Gelade 1980), cognitive (e.g., Rosch et al. 1976), and motor (e.g., Rosenbaum et al. 1983) activity. Postulation of a hierarchical structure in displacement is consistent with proposals that (a) top-down expectation can influence the perceived location of a target, and (b) bottom-up activity can modulate the remembered location of a target. Indeed, it is possible that aspects of displacement attributed to perception (e.g., Kerzel 2002c) involve such a top-down influence on perception that is based on information or expectations in memory.

20.3.2.3 Integration of levels

Given that both low- and high-level variables influence displacement, any unified account must address how to bridge low-level variables and theories to high-level variables and theories. However, acknowledgment that both low- and high-level variables can influence displacement does not entail that variables at different levels are equivalent or equal in importance (e.g., high-level expectations might drive low-level eye movements). For example, given the greater breadth of high-level theories noted earlier, any integration of low- and high-level information might be more easily accomplished at a high level. A high-level mechanism might more easily access information from high-level (e.g., expectations and beliefs regarding target motion) or low-level (e.g., pursuit eye movements) sources, and might more easily control high-level (e.g., allocation of attention) or low-level (e.g., anticipatory eye movements) processes. Also, as Hubbard (2005) noted, a high-level mechanism that biases encoding or retrieval of location would be consistent with high-level schemata, scripts, and frames that bias encoding or retrieval in other domains of memory. Any future theory of displacement should seek not merely an integrated or unified theory of displacement and representational momentum, but to integrate displacement within the larger range of cognitive processes organisms engage in on a daily basis.

20.3.2.4 Focus on function (and functionalism)

As noted earlier, a computational theory of a given process is not initially concerned with implementation of that process. Along these lines, a functionalist approach suggests that the usefulness or appropriateness of a cognitive theory of a given process does not require a given outcome (e.g., forward displacement) be instantiated by the same mechanism or implementation each time that outcome occurs; rather, what is important is the relationship between that outcome and other cognitive processes, sensory inputs, or motor outputs. Displacement for implied motion targets and continuous motion targets reflects the same relationship between inputs and outputs (i.e., memory for a moving target is displaced forward), and so displacement for both types of stimuli could reflect the same general function, even though displacement for implied motion targets and continuous motion targets might involve different forms of implementation. A unified theory of displacement could potentially be based on the idea that different examples of displacement involve the same function, rather than on findings that different examples of displacement are implemented in different ways. Also, a focus on function would shed additional light on how displacement aids the organism, what problems displacement helps solve, and would help elaborate a computational theory of displacement.

20.4 Conclusion

Memory for the final location of a moving target is often displaced in the direction of previous or anticipated target motion, and this has been called *representational momentum*. A wide range of theories and models have been proposed to account for representational momentum and related types of displacement, with some theories applying only to a specific type of stimulus or response, and other theories applying to a wider range of stimulus or response types. These theories and models cover a range of approaches, from low-level explanations based on oculomotor behavior, motion aftereffects, and perceptual adaptation, to high-level explanations based on an observer's beliefs regarding physical systems and properties of mental representation. A classification system was proposed that involves (a) internalization theories, (b) belief-based theories, (c) neointernalization theories, (d) low-level theories, and (e) network models. Consistencies and inconsistencies of each of the individual theories and models with empirical findings were noted. Neointernalization theories and network models seem consistent with more overall data than are other types of theories, but no single theory or model accounts for all the existent data on displacement or representational momentum.

The possibility of a more unified theory of displacement was addressed. Recent approaches involving multiple factors (eye movements, mode of responding, type of motion) and a proposal for a computational theory of displacement were briefly discussed, and it was suggested that any future theory of displacement (in general) and representational momentum (in particular) should be based on a consideration of function and not the different implementations that might contribute to production of displacement or representational momentum. Several considerations for future development were described including addressing (a) an increased breadth of data, (b) the existence of multiple levels of processing, (c) how low- and high-level variables and information might be combined or integrated, and (d) the function of displacement. Finally, a parallel between representational momentum and high-level schemata, scripts, and frames that bias encoding and retrieval in other domains of cognition was noted, and it was suggested that any future theory of displacement should try to integrate representational momentum and displacement within this larger range of cognitive processes.

References

Bertamini, M. (2002). Representational momentum, internalized dynamics, and perceptual adaptation. *Vis Cogn* **9**: 195–216.

Bonnet, C., Le Gall, M., & Lorenceau, J. (1984). Visual motion aftereffects: adaptation and conditioned processes. In L. Spillman & B. R. Wooten (eds.), *Sensory Experience, Adaptation, and Perception*. Hillsdale, NJ: Erlbaum.

Brehaut, J. C., & Tipper, S. P. (1996). Representational momentum and memory for luminance. *J Exp Psychol Hum Percept Perform* **22**: 480–501.

Brouwer, A. M., Franz, V. H., & Thornton, I. M. (2004). Representational momentum in perception and grasping: translating versus transforming object. *J Vis* **4**: 575–584.

Conners, F. A., Wyatt, B. S., & Dulaney, C. L. (1998). Cognitive representation of motion in individuals with mental retardation. *Am J Ment Retard* **102**: 438–450.

Cooper, L. A., & Munger, M. P. (1993). Extrapolations and remembering positions along cognitive trajectories: uses and limitations of analogies to physical momentum. In N. Eilen, R. McCarthy, & B. Brewer (eds.), *Spatial Representation: Problems in Philosophy and Psychology* (112–131). Cambridge, MA: Blackwell.

Courtney, J. R., & Hubbard, T. L. (2008). Spatial memory and explicit knowledge: an effect of instruction on representational momentum. *Q J Exp Psychol* **61**: 1778–1784.

Dawson, M. R. W. (1998). *Understanding Cognitive Science*. Cambridge, MA: Blackwell.

Desmurget, M., & Grafton, S. (2003). Feedback or feedforward control: end of a dichotomy. In S. H. Johnson-Frey (ed.), *Taking Action: Cognitive Neuroscience Perspectives on Intentional Acts* (289–338). Cambridge, MA: MIT Press.

Erlhagen, W. (2003). Internal models for visual perception. *Biol Cybern* **88**: 409–417.

Erlhagen, W., & Jancke, D. (1999). Motion waves in primary visual cortex as a neural correlate for the perception of moving objects. *Abstracts of the Society Neuroscience* **25**: 679.

Erlhagen, W., & Jancke, D. (2004). The role of action plans and other cognitive factors in motion extrapolation: a modeling study. *Vis Cogn* **11**: 315–340.

Faust, M. (1990). *Representational Momentum: A Dual Process Perspective*. Unpublished doctoral dissertation, University of Oregon, Eugene, OR.

Finke, R. A., & Freyd, J. J. (1985). Transformations of visual memory induced by implied motions of pattern elements. *J Exp Psychol Learn Mem Cogn* **11**: 780–794.

Finke, R. A., & Freyd, J. J. (1989). Mental extrapolation and cognitive penetrability: reply to Ranney and proposals for evaluative criteria. *J Exp Psychol Gen* **118**: 403–408.

Finke, R. A., Freyd, J. J., & Shyi, G. C. W. (1986). Implied velocity and acceleration induce transformations of visual memory. *J Exp Psychol Gen* **115**: 175–188.

Finke, R. A., & Shyi, G. C. W. (1988). Mental extrapolation and representational momentum for complex implied motions. *J Exp Psychol Learn Mem Cogn* **14**: 112–120.

Freyd, J. J. (1983). The mental representation of movement when static stimuli are viewed. *Percept Psychophys* **33**: 575–581.

Freyd, J. J. (1987). Dynamic mental representations. *Psychol Rev* **94**: 427–438.

Freyd, J. J. (1992). Dynamic representations guiding adaptive behavior. In F. Macar, V. Pouthas, & W. J. Friedman (eds.), *Time, Action, and Cognition: Towards Bridging the Gap* (309–323). Dordrecht: Kluwer Academic Publishers.

Freyd, J. J. (1993). Five hunches about perceptual processes and dynamic representations. In D. Meyer & S. Kornblum (eds.), *Attention and Performance XIV: Synergies in Experimental Psychology, Artificial Intelligence, and Cognitive Neuroscience* (99–119). Cambridge, MA: MIT Press.

Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *J Exp Psychol Learn Mem Cogn* **10**: 126–132.

Freyd, J. J., & Finke, R. A. (1985). A velocity effect for representational momentum. *Bull Psychon Soc* **23**: 443–446.

Freyd, J. J., & Johnson, J. Q. (1987). Probing the time course of representational momentum. *J Exp Psychol Learn Mem Cogn* **13**: 259–269.

Freyd, J. J., & Jones, K. T. (1994). Representational momentum for a spiral path. *J Exp Psychol Learn Mem Cogn* **20**: 968–976.

Freyd, J. J., Kelly, M. H., & DeKay, M. L. (1990). Representational momentum in memory for pitch. *J Exp Psychol Learn Mem Cogn* **16**: 1107–1117.

Freyd, J. J., Pantzer, T. M., & Cheng, J. L. (1988). Representing statics as forces in equilibrium. *J Exp Psychol Gen* **117**: 395–407.

Futterweit, L. R., & Beilin, H. (1994). Recognition memory for movement in photographs: a developmental study. *J Exp Child Psychol* **57**: 163–179.

Getzmann, S., Lewald, J., & Guski, R. (2004). Representational momentum in spatial hearing. *Perception* **33**: 591–599.

Halpern, A. R., & Kelly, M. H. (1993). Memory biases in left versus right implied motion. *J Exp Psychol Learn Mem Cogn* **19**: 471–484.

Hayes, A. E., & Freyd, J. J. (2002). Representational momentum when attention is divided. *Vis Cogn* **9**: 8–27.

Hubbard, T. L. (1990). Cognitive representation of linear motion: possible direction and gravity effects in judged displacement. *Memory & Cognition* **18**: 299–309.

Hubbard, T. L. (1993a). Auditory representational momentum: musical schemata and modularity. *Bull Psychon Soc* **31**: 201–204.

Hubbard, T. L. (1993b). The effects of context on visual representational momentum. *Memory & Cognition* **21**: 103–114.

Hubbard, T. L. (1994). Judged displacement: a modular process? *Am J Psychol* **107**: 359–373.

Hubbard, T. L. (1995a). Cognitive representation of motion: evidence for representational friction and gravity analogues. *J Exp Psychol Learn Mem Cogn* **21**: 241–254.

Hubbard, T. L. (1995b). Environmental invariants in the representation of motion: implied dynamics and representational momentum, gravity, friction, and centripetal force. *Psychon Bull Rev* **2**: 322–338.

Hubbard, T. L. (1996). Representational momentum, centripetal force, and curvilinear impetus. *J Exp Psychol Learn Mem Cogn* **22**: 1049–1060.

Hubbard, T. L. (1997). Target size and displacement along the axis of implied gravitational attraction: effects of implied weight and evidence of representational gravity. *J Exp Psychol Learn Mem Cogn* **23**: 1484–1493.

Hubbard, T. L. (1998a). Representational momentum and other displacements in memory as evidence for nonconscious knowledge of physical principles. In S. Hameroff, A. Kaszniak, & A. Scott (eds.), *Towards a Science of Consciousness II: The Second Tucson Discussions and Debates* (505–512). Cambridge, MA: MIT Press.

Hubbard, T. L. (1998b). Some effects of representational friction, target size, and memory averaging on memory for vertically moving targets. *Can J Exp Psychol* **52**: 44–49.

Hubbard, T. L. (1999). How consequences of physical principles influence mental representation: the environmental invariants hypothesis. In P. R. Killeen & W. R. Uttal (eds.), *Fechner Day 99: The End of 20th Century Psychophysics. Proceedings of the 15th Annual Meeting of the International Society for Psychophysics* (274–279). Tempe, AZ: The International Society for Psychophysics.

Hubbard, T. L. (2004). The perception of causality: insights from Michotte's launching effect, naive impetus theory, and representational momentum. In A. M. Oliveira, M.P. Teixeira, G. F. Borges, & M. J. Ferro (eds.), *Fechner Day 2004* (116–121). Coimbra, Portugal: The International Society for Psychophysics.

Hubbard, T. L. (2005). Representational momentum and related displacements in spatial memory: a review of the findings. *Psychon Bull Rev* **12**: 822–851.

Hubbard, T. L. (2006a). Bridging the gap: possible roles and contributions of representational momentum. *Psicologica* **27**: 1–34.

Hubbard, T. L. (2006b). Computational theory and cognition in representational momentum and related types of displacement: a reply to Kerzel. *Psychon Bull Rev* **13**: 174–177.

Hubbard, T. L., & Bharucha, J. J. (1988). Judged displacement in apparent vertical and horizontal motion. *Percept Psychophys* **44**: 211–221.

Hubbard, T. L., Blessum, J. A., & Ruppel, S. E. (2001). Representational momentum and Michotte's (1946/1963) "Launching Effect" paradigm. *J Exp Psychol Learn Mem Cogn* **27**: 294–301.

Hubbard, T. L., & Courtney, J. R. (2006). Evidence for a separation of perceptual and cognitive dynamics. In L. Albertazzi (ed.), *Visual Depictive Thought* (71–97). New York: Benjamins Publishing Company.

Hubbard, T. L., & Favretto, A. (2003). Explorations of Michotte's "Tool Effect": evidence from representational momentum. *Psychol Res* **67**: 134–152.

Hubbard, T. L., Kumar, A. M., & Carp, C. L. (2009). Effects of spatial cueing on representational momentum. *J Exp Psychol Learn Mem Cogn*. **35**: 666–677.

Hubbard, T. L., Matzenbacher, D. L., & Davis, S. E. (1999). Representational momentum in children: dynamic information and analogue representation. *Percept Mot Skills* **88**: 910–916.

Hubbard, T. L., & Ruppel, S. E. (1999). Representational momentum and landmark attraction effects. *Can J Exp Psychol* **53**: 242–256.

Hubbard, T. L., & Ruppel, S. E. (2002). A possible role of naive impetus in Michotte's "Launching Effect:" evidence from representational momentum. *Vis Cogn* **9**: 153–176.

Jarrett, C. B., Phillips, M., Parker, A., & Senior, C. (2002). Implicit motion perception in schizotypy and schizophrenia: a representational momentum study. *Cogn Neuropsychiatry* **7**: 1–14.

Johnston, H., & Jones, M. R. (2006). Higher-order pattern structure influences auditory representational momentum. *J Exp Psychol Hum Percept Perform* **32**: 2–17.

Joordens, S., Spalek, T. M., Razmy, S., & van Duijn, M. (2004). A clockwork orange: compensation opposing momentum in memory for location. *Memory & Cognition* **32**: 39–50.

Jordan, J. S. (1998). Recasting Dewey's critique of the reflex-arc concept via a theory of anticipatory consciousness: implications for theories of perception. *New Ideas Psychol* **16**: 165–187.

Jordan, J. S., & Hunsinger, M. (2008). Learned patterns of action-effect extrapolation contribute to the spatial displacement of continuously moving stimuli. *J Exp Psychol Hum Percept Perform* **34**(1): 113–124.

Jordan, J. S., & Knoblich, G. (2004). Spatial perception and control. *Psychon Bull Rev* **11**: 54–59.

Jordan, J. S., Stork, S., Knuf, L., Kerzel, D., & Müsseler, J. (2002). Action planning affects spatial localization. In W. Prinz & B. Hommel (eds.), *Attention and Performance XIX: Common Mechanisms in Perception and Action*. (158–176). New York: Oxford University Press.

Kaiser, M. K., Proffitt, D. R., & Anderson, K. (1985). Judgments of natural and anomalous trajectories in the presence and absence of motion. *J Exp Psychol Learn Mem Cognition* **11**: 795–803.

Kaiser, M. K., Proffitt, D. R., Whelan, S. M., & Hecht, H. (1992). Influence of animation on dynamical judgments. *J Exp Psychol Hum Percept Perform* **18**: 669–690.

Kelly, M. H., & Freyd, J. J. (1987). Explorations of representational momentum. *Cogn Psychol* **19**: 369–401.

Kerzel, D. (2000). Eye movements and visible persistence explain the mislocalization of the final position of a moving target. *Vision Res* **40**: 3703–3715.

Kerzel, D. (2002a). A matter of design: no representational momentum without predictability. *Vis Cogn* **9**: 66–80.

Kerzel, D. (2002b). Attention shifts and memory averaging. *Q J Exp Psychol* **55**(A): 425–443.

Kerzel, D. (2002c). The locus of "memory displacement" is at least partially perceptual: effects of velocity, expectation, friction, memory averaging, and weight. *Percept Psychophys* **64**: 680–692.

Kerzel, D. (2003a). Attention maintains mental extrapolation of target position: irrelevant distractors eliminate forward displacement after implied motion. *Cognition* **88**: 109–131.

Kerzel, D. (2003b). Centripetal force draws the eyes, not memory of the target, toward the center. *J Exp Psychol Learn Mem Cogn* **29**: 458–466.

Kerzel, D. (2003c). Mental extrapolation of target position is strongest with weak motion signals and motor responses. *Vision Res* **43**: 2623–2635.

Kerzel, D. (2005). Representational momentum beyond internalized physics. *Curr Dir Psychol Sci* **14**: 180–184.

Kerzel, D. (2006). Why eye movements and perceptual factors have to be controlled in studies on "representational momentum." *Psychon Bull Rev* **13**: 166–173.

Kerzel, D., & Gegenfurtner, K. R. (2003). Neuronal processing delays are compensated in the sensorimotor branch of the visual system. *Curr Biol* **13**: 1975–1978.

Kerzel, D., Jordan, J. S., & Müsseler, J. (2001). The role of perception in the mislocalization of the final position of a moving target. *J Exp Psychol Hum Percept Perform* **27**: 829–840.

Kozhevnikov, M., & Hegarty, M. (2001). Impetus beliefs as default heuristics: dissociation between explicit and implicit knowledge about motion. *Psychon Bull Rev* **8**: 439–453.

Marr, D. (1982). *Vision*. New York: W. H. Freeman and Company.

McCloskey, M. (1983). Naive theories of motion. In D. Gentner & A. L. Stevens (eds.), *Mental Models* (299–324). Hillsdale, NJ: Erlbaum.

McCloskey, M., & Kohl, D. (1983). Naive physics: the curvilinear impetus principle and its role in interactions with moving objects. *J Exp Psychol Learn Mem Cogn* **9**: 146–156.

Michotte, A. (1963). *The Perception of Causality* (T. R. Miles & E. Miles, Trans.). New York: Basic Books (original work published 1946).

Motes, M. A., Hubbard, T. L., Courtney, J. R., & Rypma, B. (2008). A principal components analysis of dynamic spatial memory biases. *J Exp Psychol Learn Mem Cogn* **34**: 1076–1083.

Munger, M. P., & Minchew, J. H. (2002). Parallels between remembering and predicting an object's location. *Vis Cogn* **9**: 177–194.

Munger, M. P., & Owens, T. R. (2004). Representational momentum and the flash-lag effect. *Vis Cogn* **11**: 81–103.

Munger, M. P., Solberg, J. L., & Horrocks, K. K. (1999). On the relation between mental rotation and representational momentum. *J Exp Psychol Learn Mem Cogn* **25**: 1557–1568.

Munger, M. P., Solberg, J. L., Horrocks, K. K., & Preston, A. S. (1999). Representational momentum for rotations in depth: effects of shading and axis. *J Exp Psychol Learn Mem Cogn* **25**: 157–171.

Müsseler, J., & Aschersleben, G. (1998). Localizing the first position of a moving stimulus: the Fröhlich effect and an attention-shifting explanation, *Percept Psychophys* **60**: 683–695.

Müsseler, J., Stork, S., & Kerzel, D. (2002). Comparing mislocalizations with moving stimuli: the Fröhlich effect, the flash-lag, and representational momentum. *Vis Cogn* **9**: 120–138.

Müsseler, J., van der Heijden, A. H. C., Mahmud, S. H., Deubel, H., & Ertsey, S. (1999). Relative mislocalization of briefly presented stimuli in the retinal periphery. *Percept Psychophys* **61**: 1646–1661.

Nijhawan, R. (2002). Neural delays, visual motion and flash-lag effect. *Trends Cogn Sci* **6**: 387–393.

Poljansek, A. (2002). The effect of motion acceleration on displacement of continuous and staircase motion in the frontoparallel plane. *Psiholoska obzorja/Horizons of Psychology* **11**: 7–21.

Raymond, J. E., O'Donnell, H. L., & Tipper, S. P. (1998). Successive episodes produce direction contrast effects in motion perception. *Vision Res* **38**: 579–590.

Reed, C. L., & Vinson, N. G. (1996). Conceptual effects on representational momentum. *J Exp Psychol Hum Percept Perform* **22**: 839–850.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cogn Psychol* **8**: 382–439.

Rosenbaum, D. A., Kenny, S., & Derr, M. A. (1983). Hierarchical control of rapid motor sequences. *J Exp Psychol Hum Percept Perform* **9**: 86–102.

Ruppel, S. E., Fleming, C. N., & Hubbard, T. L. (2009). Representational momentum is not (totally) impervious to error feedback. *Can J Exp Psychol* **63**: 49–58.

Senior, C., Barnes, J., & David, A. S. (2001). Mental imagery increases representational momentum: preliminary findings. *J Ment Imagery* **25**: 177–184.

Shepard, R. N. (1975). Form, formation, and transformation of internal representations. In R. L. Solso (ed.), *Information Processing and Cognition: The Loyola Symposium* (87–122). Hillsdale, NJ: Erlbaum.

Shepard, R. N. (1981). Psychophysical complementarity. In M. Kubovy & J. R. Pomerantz (eds.), *Perceptual Organization* (279–341). Hillsdale, NJ: Erlbaum.

Shepard, R. N., & Chipman, S. (1970). Second-order isomorphism of internal representations: shapes of states. *Cogn Psychol* **1**: 1–17.

Thornton, I. M., & Hayes, A. E. (2004). Anticipating action in complex scenes. *Vis Cogn* **11**: 341–370.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cogn Psychol* **12**: 97–136.

Verfaillie, K., & d'Ydewalle, G. (1991). Representational momentum and event course anticipation in the perception of implied periodical motions. *J Exp Psychol Learn Mem Cogn* **17**: 302–313.

Vinson, N. G., & Reed, C. R. (2002). Sources of object-specific effect in representational momentum. *Vis Cogn* **9**: 41–65.

Whitney, D., & Cavanagh, P. (2002). Surrounding motion affects the perceived locations of moving stimuli. *Vis Cogn* **9**: 139–152.

21

Conceptual influence on the flash-lag effect and representational momentum

MASAYOSHI NAGAI, MUTSUMI SUGANUMA, ROMI NIJHAWAN,
JENNIFER J. FREYD, GEOFFREY MILLER, AND KATSUMI WATANABE

21.1 Introduction

When judging the position of a *moving* object, human observers do not perceive and memorize the moving object's correct position. There are two known phenomena in judged position errors of a moving object, representational momentum (RM) and the flash-lag effect (FLE), both of which we consider here.

RM was originally reported by Freyd and Finke (1984). Freyd and colleagues displayed a series of still frames to imply the rotation of a rectangle (e.g., Freyd & Finke 1984, 1985; Freyd & Johnson 1987). Observers saw three views of a rectangle at different rotations about its center, with 250 msec display duration with 250 msec interstimulus interval. The fourth rectangle was presented as a probe 250 msec after the third frame presentation. The rotation of the probe was selected from possible positions symmetrically distributed around the actual third position of the rectangle. Observers were asked whether the rectangle in the third frame (the last frame of the motion sequence) was the same orientation as the probe. The results showed that their memory for the third orientation tended to be shifted in the direction of rotation. In other words, the orientation of the probe rectangle had to be rotated slightly further to be judged as being in the same position as the third rectangle. To account for the forward shift of the final position of a stimulus undergoing implied motion, some authors postulate that the dynamics of the representational system follow physical laws, such as momentum (representational momentum; Finke & Freyd 1985; Finke et al. 1986; Freyd 1987; Finke & Shyi 1988). RM is a robust effect as observed with smooth object motion and in pointing at the final position of a moving object (e.g., Hubbard & Bharucha 1988). Several variables influence RM (for review, Hubbard 1995b). RM increases with the velocity (e.g., Freyd & Finke 1985; Hubbard & Bharucha 1988; Nagai & Saiki 2005) and acceleration of the moving target (Finke et al. 1986), pointing to the similarity between RM and physical momentum. Hubbard and others demonstrated that RM may reflect physical principles. For example, RM increases downward, in the direction of gravity (Hubbard & Bharucha 1988; 1995a, 1997; Reed & Vinson 1996; Hubbard & Ruppel 1999; Nagai et al. 2002; Vinson & Reed 2002), whereas implied friction between a moving stimulus and an adjoining surface reduces RM (Hubbard 1995a; see also Nagai & Yagi 2001). Moreover, RM is influenced by real-world knowledge of the typical motions of familiar

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

objects (Freyd & Miller 1992; Reed & Vinson 1996; Vinson & Reed 2002; but see Halpern & Kelly 1993; Nagai & Yagi 2001), the future or expected trajectory of a target (Hubbard & Bharucha 1988; Verfaillie & d'Ydewalle 1991), and visual attention (Hayes & Freyd 2002). Finally, Nagai and Saiki (2005) found that RM is elicited in the physical/actual direction and actual speed of an object's motion but not in its perceived direction and speed when the physical/actual versus perceived motions are different. Thus, a variety of factors affect RM: from low-level, perceptual factors such as a moving object's speed and acceleration to higher-level, cognitive factors such as expectation, attention, and each object's typical motion in the real world.

When a brief flash is presented adjacent to a continuously moving stimulus, the flash appears to lag behind the moving object. This flash-lag effect, FLE (Fröhlich 1923; Metzger 1932; Mackay 1958; Nijhawan 1994), is also robust and has been replicated in various stimulus configurations (Baldo & Klein 1995; Khurana & Nijhawan 1995; Nijhawan 1997; Kirschfeld & Kammer 1999; Brenner & Smeets 2000; Eagleman & Sejnowski 2000a; Khurana et al. 2000; Watanabe et al. 2001; Watanabe et al. 2003; Watanabe 2004). Several factors are known to influence the FLE. The FLE increases as the luminance of the moving object is increased and decreases as the luminance of the flash is increased (Purushothaman et al. 1998). It increases as the retinal eccentricity of the flash is increased (Baldo & Klein 1995). Recently, Anstis (2007) found FLE occurs in the physical, not the subjective, direction of rotation. In addition to these low-level stimulus factors, perceptual grouping causes a large modulation of FLE magnitude (Watanabe et al. 2001; Watanabe 2004). FLE is reduced when the observer knows when and where the next flash is to occur (Brenner & Smeets 2000; Eagleman & Sejnowski 2000b; also see Nagai et al. 2000). Lastly, there is an ongoing debate about whether FLE is affected by attention (Khurana et al. 2000; Baldo et al. 2002). Thus, it is unclear whether the factors influencing FLE are limited to low-level or perceptual ones or also include higher cognitive factors similar to the factors affecting RM. Previously, it was shown that forward motion of objects that have a normal motion direction (e.g., animals that typically move headfirst) causes a larger RM effect than backward motion (Freyd & Miller 1992; Reed & Vinson 1996; Vinson & Reed 2002). Here we tested whether knowledge of an object's typical motion in the world influences FLE and directly compared it with that of RM. If such knowledge influences FLE, then forward motion would cause larger FLE than backward motion.

21.2 A flash-lag effect experiment

Here we tested the influence of object-typical motions on FLE. There were three different conditions: forward, backward, and stationary conditions (i.e., control condition). In the forward condition, the picture of a car moved forward, whereas in the backward condition it moved backward. In the control condition, the picture did not move. If the influence of object-typical motions exists in the FLE, then FLE in the forward condition should be larger than in the backward condition, as shown in RM studies (Freyd & Miller 1992; Reed & Vinson 1996; Vinson & Reed 2002).

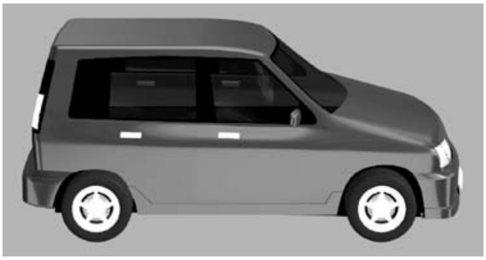


Fig. 21.1 An example of picture stimuli used in the flash-lag experiment.

21.2.1 Method

21.2.1.1 Observers

Six adults (age range 20–27 years; mean = 21.6 years) served as participants. They all had normal or corrected-to-normal vision.

21.2.1.2 Apparatus

The stimuli were produced using an Apple Power Macintosh G4 computer (with Mac OS 9.2) and were displayed on a CRT monitor (Sony 21-inch color monitor, refresh rate 75 Hz) in a dimly lit room. The viewing distance was 57 cm, and a chin rest was employed to maintain constant viewing distance.

21.2.1.3 Stimuli

Stimuli were presented on a gray background. The moving object was a picture of a car, as shown in Fig. 21.1. The picture size was 1.2-deg wide and 0.7-deg high. A white fixation cross was presented at the center of the screen, and a white dot (0.08 deg in diameter) was presented as a flash probe. The object moved horizontally by 5 pixels per frame (corresponding to 7.5 deg/s), with the movement trajectory above the fixation point. The distance in the vertical dimension from the fixation point to the bottom of the car was approximately 0.2 deg. The flash was presented for one frame ($\approx$ 13.3 msec) 1.4 deg above the fixation point. The vertical distance between the roof of the car and the position of the flash was approximately 0.5 deg. The car picture appeared 5.6 deg left or right from the fixation point, moved toward the center of the screen, and went through to the opposite side of the screen. The duration of the complete motion sequence was 1533.3 msec (115 frames). The flash and the fixation point were always aligned on the vertical axis. However, the relative horizontal position between the flash and the car picture varied from trial to trial according to lag condition.

21.2.1.4 Procedure

The observer's task was to decide whether the flash appeared to the left or right relative to the "center" of the car picture (two-alternative forced choice (2AFC)). Observers were strictly instructed to fixate on a fixation cross while it was presented on the screen.

The experimental design was as follows. For the factor of object motion, there were three types of motion. In the conditions in which the car picture moved (forward, backward conditions), the picture was initially presented 5.6 deg left or right of fixation. When the car picture appeared on the left side, it moved toward the right, and vice versa. The car picture moved at a constant speed of 7.5 deg/sec to the opposite side of the fixation point. When the car picture reached 5.6 deg to the opposite side of the fixation point, it disappeared. In the condition in which the car did not move (stationary condition), the picture was presented above the fixation point (with spatial shift according to lag condition) for 1533.3 msec. For the factor of object orientation, there were two types of object orientation, the picture of the car facing left or facing right.

As for the factor of horizontal lag between the flash and the picture of the car, we used seven different lags: −0.6, −0.4, −0.2, 0, 0.2, 0.4, 0.6 deg (−6, −4, −2, 0, 2, 4, and 6 temporal frames in the time domain). We expediently defined the horizontal midpoint of the car picture as its center. When the lag was 0 deg, the flash appeared above the fixation cross when the center of the car was just above the fixation. In other lag settings, when the lag was −0.6 deg, the flash was presented six temporal frames before the center of the car picture reached a point directly above the fixation point. Observers performed 672 trials in total: 3 (leftward motion, rightward motion, or stationary) × 2 (facing left or right) × 7 (different lag settings) × 16 (repetitions of each condition).

At the beginning of each trial, the fixation cross was presented at the center of the screen. The car picture appeared on the screen 500 msec after the fixation cross appearance. In all conditions the car picture was presented for 1533.3 msec. Observers made their responses by pressing the left or right arrow key, after which the fixation cross disappeared. No feedback was given to the observers. The next trial was then presented with a 1-sec intertrial interval. A short break was given to the observers every sixty trials (approximately once every 4 min), which the observers used to take a short rest if needed. Prior to the experiment, observers completed a practice session consisting of thirty trials. The entire experiment time was about 60 min, including instruction and practice session.

21.2.2 Result and discussion

Figure 21.2 shows the averaged FLE in each condition for six observers in the FLE experiment. The data were collapsed across motion direction because there was no difference between leftward and rightward motion, and recombined into forward (leftward-facing left trials and rightward-facing right trials) or backward motion (leftward-facing right trials and rightward-facing left trials). Thus, in this experiment, the *actual amount* of flash lag was defined as the difference between the stationary condition (which was used to estimate each observer's "subjective" center of the car picture) and the forward motion condition, or between the stationary condition and the backward motion condition. For each observer, we derived a psychometric function and calculated the 50% probability of judging "left" or "right" for each of the conditions. We observed significant differences between conditions ($F(2, 10) = 14.5$, $\text{MSE} = 20.7$, $p < .01$). The FLE was significantly smaller for the

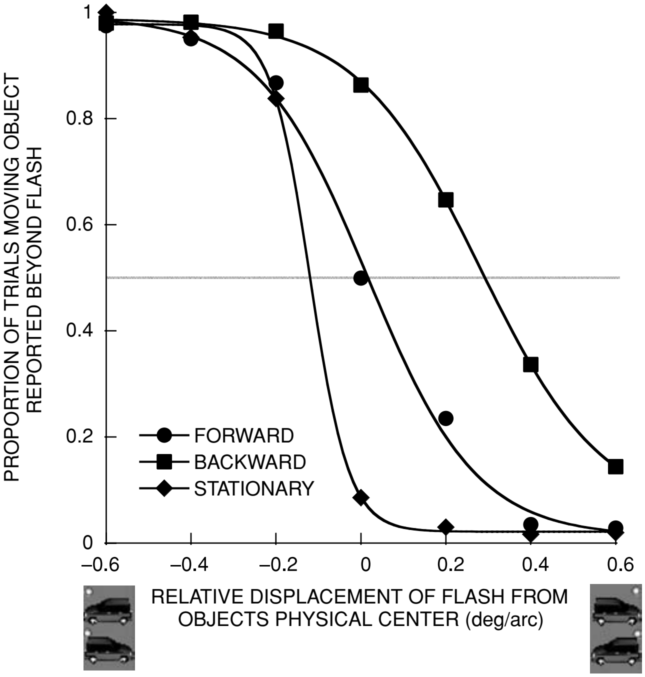


Fig. 21.2 Result of flash-lag experiment.

stationary condition than for the other two conditions ($p < 0.01$ for both forward and backward conditions), which indicates that FLE occurred in this display. Then, we took the difference in the probability judgment between the forward motion condition and the stationary condition as representing the actual amount of FLE in forward motion, and the difference in the probability judgment between the backward motion condition and the stationary motion condition as representing the actual amount of FLE in backward motion. A separate statistical analysis revealed that the difference in FLE between the forward motion condition and the backward motion condition was significant [$t(5) = 5.5, p < 0.01$]. A larger FLE occurred for the backward motion condition than for the forward motion condition.

This result, that backward motion produced larger FLE than forward motion, was surprising, because it was opposite to the findings in RM (Freyd & Miller 1992; Reed & Vinson 1996; Vinson & Reed 2002). We replicated this result with a different stimulus (a picture of a fish shown in Fig. 21.3), and thus the effect of forward/backward motions on FLE (i.e., the opposite effect shown in RM) seems robust. It is worth considering the quality of motion (discrete or smooth motion) in discussing knowledge-based effects. In RM, with discrete motion, the effect of typical motion was consistently found (Reed & Vinson 1996; Vinson & Reed 2002), but it was not consistently found with smooth motion (Freyd & Miller 1992; Nagai & Yagi 2001). These findings suggested that discrete motion of an object was better to show a typical motion effect in RM. However, in the present experiment, we followed a standard experimental procedure of FLE studies for better comparison of FLE with RM.

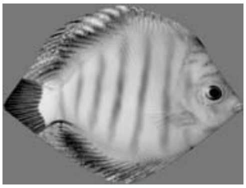


Fig. 21.3 Example of stimuli used in preliminary test.

Before we consider why the effect in FLE was opposite to that in RM, it is necessary to confirm the typical motion effect in RM with smooth motion.

21.3 Representational momentum experiment with left or right judgment

As mentioned in the introduction, previous studies of RM have shown that conceptual knowledge of objects' and animals' typical motions in the real world influenced the magnitude of RM (Freyd & Miller 1992; Reed & Vinson 1996; Vinson & Reed 2002; but see Halpern & Kelly 1993[1]; Nagai & Yagi 2001). However, the typical motion effect was not consistently observed with smooth motion: for example, Freyd and Miller (1992) showed the effect, but Nagai and Yagi (2001) did not. Here we tried to replicate the typical motion effect with smooth motion. A standard RM paradigm employs the same–different judgment between the final position of the moving stimulus and the position of the subsequently presented probe. However, in this experiment, the left or right judgment as in the FLE experiment was used to allow a more direct comparison between the typical motion effects on FLE and RM.

21.3.1 Method

21.3.1.1 Observers

Six adults (ranged 20–27 years; mean = 22.3 years) served as participants. They all had normal or corrected-to-normal vision. Three of those participants had participated in the flash-lag experiment.

21.3.1.2 Stimuli

The stimulus used for the moving (or staying) object was the same picture of a car as used in the FLE experiment. In this RM experiment, no fixation point or flash was presented because keeping eyes at the fixation point reduces the magnitude of RM (Kerzel 2000; Nagai & Saiki 2006), and elements other than a moving object bias the judged final position of

---

[1] In Halpern and Kelly (1993), only *forward* discrete motions of a fox, a motorcycle, a rhinoceros, a truck, and a ball were used, and they did not show consistent effect of typical speeds in the real world like a truck moves faster than a fox. In contrast, most of the studies used *forward and backward* discrete motions of real-world objects and showed different magnitude of RM for these two motions (Freyd & Miller 1992; Reed & Vinson 1996; Vinson & Reed 2002). These suggest that comparing forward and backward motions was a sensitive way to show typical motion effects and the relative difference in typical speed between objects could not affect RM. In the present study we used the sensitive method to examine the typical motion effect.

the object to the other element (Hubbard & Ruppel 1999). The car picture appeared at the same position (5.6 deg left or right from the horizontal center of the display) and moved with the same speed (or remained still as in FLE stationary condition).

The disappearing location of the object was randomly set from trial to trial, in the range from -1.0 to $+1.0$ deg from the exact center of the screen. In the forward and backward motion conditions, the motion duration differed according to where the object disappeared. When the object disappeared at the center of the screen, motion duration was 746.48 msec (ranged from 613.1 msec to 879.8 msec). In the stationary condition, the car picture was presented near the center of the screen (at the center with spatial jitter ranged from -1.0 to $+1.0$ deg) for 746.48 msec. After the car picture disappeared, observers were shown a blank display (entirely gray) for 1 sec. After this retention interval, the car picture was shown again as a probe, with a spatial displacement according to condition. This probe picture remained on the screen until participants' response.

21.3.1.3 Procedure

The observer's task was to decide whether the probe appeared to the left, or right, relative to the car picture in the final frame (two-alternative forced-choice). Observers were not informed whether they should track the moving stimulus or not, as in the standard experiment paradigm of RM.

The experimental design was similar to that of the FLE experiment except for the probe presentation. There were seven different probe positions (-0.6, -0.4, -0.2, 0, $+0.2$, $+0.4$, $+0.6$ deg). Positive values indicate that the probe was shifted in the direction of the picture motion from the final position of the car picture. Observers performed 672 trials in total: 3 (leftward motion, rightward motion, or stationary) × 2 (facing left or right) × 7 (probe positions) × 16 (repetitions per condition).

The beginning of each trial was indicated with a short beep sound. The car picture appeared on the screen 500 msec after the beep. Observers made their response by pressing the left or right arrow key. No feedback was given to the observers. They went on to the next trial, after 1 sec of intertrial interval. A short break was inserted every sixty trials (approximately once every 5–6 min), and observers took a rest if they needed to. Prior to the experiment, observers completed a practice session consisting of thirty trials. The entire experiment time was about 70 min including instruction and practice session.

21.3.2 Result and discussion

Figure 21.4 shows the averaged RM shift in each condition for six observers. The shift was defined as judged error relative to the final position. Collected data were collapsed across motion direction (there was no difference between leftward and rightward motion) and recombined as forward (leftward-facing left trials and rightward-facing right trials) or backward motion (leftward-facing right trials and rightward-facing left trials).

The overall results resemble those of the previous FLE experiment. The magnitude of RM was nonsignificantly larger for backward motion than for forward motion. However, we found no significant difference from the zero baseline for any conditions, which indicates

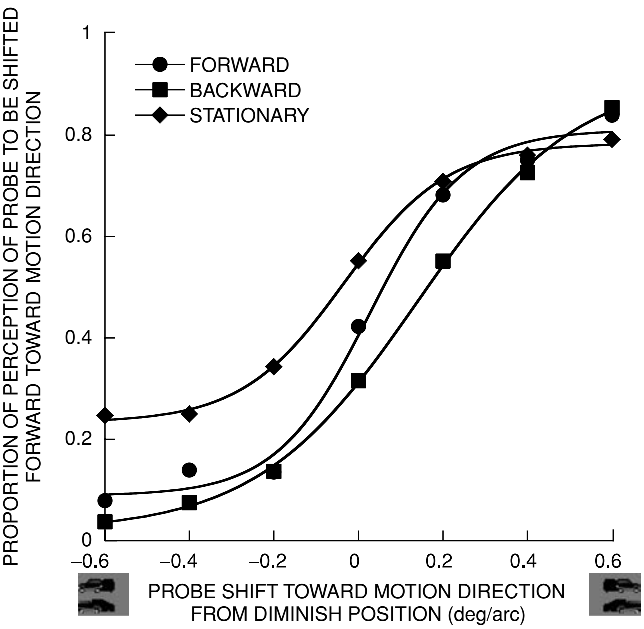


Fig. 21.4 Result for RM experiment, when judging whether the probe was shifted left or right from final frame.

that no significant RM shift occurred in this experiment. This might be due to the task we used. In this experiment, to make a direct comparison between the FLE and RM experiments, we asked observers to decide whether the probe shifted left or right. In the next experiment, we used a same/different judgment paradigm, which is the method more widely used in RM studies (e.g., Freyd & Finke 1984).

21.4 Representational momentum experiment with same/different judgment

In the previous RM experiment with the left/right judgment, the results showed no RM effects. In this experiment, each observer performed the same/different judgment on the position of the probe, but the other stimuli settings and procedures were kept identical to those in the previous experiment. This procedure was expected to increase the RM effect and to show the typical motion effect.

21.4.1 Method

21.4.1.1 Observers

Nine adults (ranged 19–31 years; mean = 22.2 years) served as participants. None had participated in the previous two experiments. They all had normal or corrected-to-normal vision.

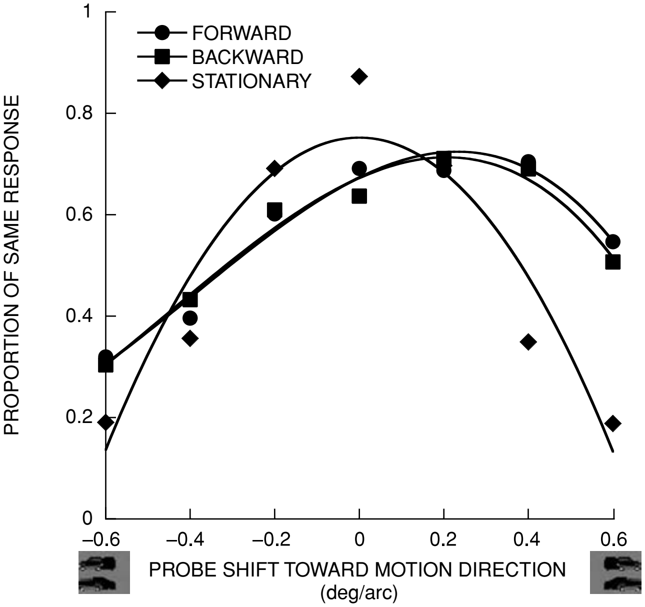


Fig. 21.5 Same response proportion and quadratic fit.

21.4.2 Result and discussion

Figure 21.5 shows the averaged proportion of the same response in each condition for nine observers in this second RM experiment with the same/different judgment. For each observer's data, we calculated a weighted mean for each condition. Figure 21.6 shows the average weighted mean and standard error for each condition. Statistical analyses revealed that the weighted mean in the forward condition was significantly larger than zero ($t(8) = 1.9$, $p < .05$) but was not in the backward or stationary conditions.

With the same/different judgment, we observed significant RM and the expected effect of forward versus backward motions. This suggests that the typical motion effect shown in previous discrete-motion RM studies (Freyd & Miller 1992; Reed & Vinson 1996; Vinson & Reed 2002) can also be observed with smooth motion, although the magnitude of it was small (but see Nagai & Yagi 2001).

21.5 General discussion

These studies aimed to directly compare the typical motion effects in FLE and RM. The first experiment showed that FLE was larger for a car's backward motion than for its forward motion; thus FLE shows a reversed-typical motion effect. This was the robust effect because it was replicated also with the motion of a biological object (e.g., a fish). Although it was the opposite to the typical motion effect found previously for RM, this study consistently showed the influence of higher cognitive knowledge about objects' typical motions on FLE.

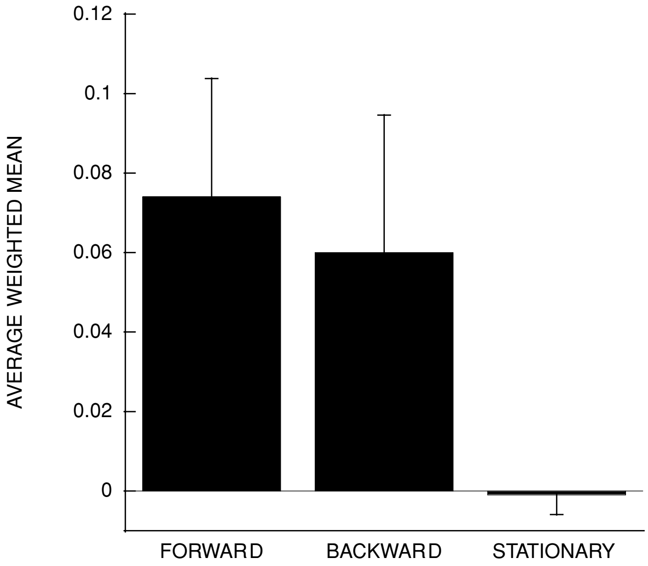


Fig. 21.6 Averaged weighted mean and standard error for each condition.

The second experiment did not show any RM for the left/right judgment task. The third experiment used the same/different judgment and replicated RM and typical motion effects that had been consistently observed before with discrete stimulus presentation (Reed & Vinson 1996; Vinson & Reed 2002). Although we used different tasks in FLE and RM, we presented the same motion stimuli in both FLE and RM experiments. Thus, the present study provides the first basis for comparing and discussing the effect of cognitive knowledge on the two motion-based spatial errors.

Here we attempt to explain the discrepant results between FLE and RM with two different typical motion effects: a "perceived speed" while a moving object is *actually* presented versus an "internal model speed signal" after the object has disappeared. In the real world, cars move forward most of time when they move, and their backward motions are much less frequent. Thus, humans could have an internal model for each type of object's typical motions: for example, a car mostly moves forward and only sometimes backward. In the case of car motion, the subjective speed estimate from the internal model would be larger for forward motion and smaller for backward motion. However, because objects in both forward and backward conditions move with the same actual speed on the CRT display, the difference between incoming sensory input and the output of the internal model would be smaller for forward motion and larger for backward motion. If perceptual speed is modified by the magnitude of such a difference (or subjective prediction error), then the perceived speed of backward motion would be larger than forward motion, and thus FLE would be larger for backward motion.

In the RM experiment paradigm, however, this perceived speed modulation by subjective prediction error could not occur, and only the internal model speed could influence the magnitude of RM. The critical difference between FLE and RM stimulus presentations was whether the car's motion was still presented at the time of the "judgment marker." In the FLE experiment, the car motion was still presented at the time of the flash, and observers could access and use this sensory input signal to make their judgment. Thus, the perceived speed modulation (based on the subjective prediction error) influenced the magnitude of FLE. However, in the RM paradigm, no moving car was presented at the time of probe presentation, which meant there was no sensory input signal of motion at the time of judgment. Thus, the perceived speed modulation did not occur in the RM paradigm, and we did not get the result that a car's backward motion produced larger RM than its forward motion. Instead, we suggest that the "internal model" on an objects' typical motions influenced RM. After the moving car disappeared, observers could not use the sensory input of the motion signal but used only the output speed signal of the internal model of the object's typical motions (e.g., a car moves faster for its forward than backward motion). Therefore, larger RM for the forward motion was observed than for the backward motion. This is the typical motion effect as found in previous studies (Freyd & Miller 1992; Reed & Vinson 1996; Vinson & Reed 2002).

This internal model of an object's typical motions could also work during the blank screen intervals during discrete motion presentation. For example, in Vinson and Reed (2002) the picture of the object was displayed for 250 msec and followed by a 250 msec blank interstimulus interval (ISI). The motion sequence consisted of four frames of object presentations interleaved with ISIs. In this discrete motion case, the internal model's influence could be applied during each ISI, which would reduce the greater perceived speed for backward motion, and which would increase the lower perceived speed for forward motion, to yield the typical motion effect (i.e., the forward motion of a car or fish producing larger RM than their backward motion).

In contrast, this study, Nagai and Yagi (2001) and Freyd and Miller (1992) used smooth motion of the object (i.e., no ISI), thus some of them yielding the small typical motion effect (this study) or no such effect (Nagai & Yagi 2001) in RM. In the case of smooth motion, the typical motion influence by the internal model could not occur during its motion presentation because there were no blank ISIs. Thus, this internal model influence might be overcome by the perceived speed modulation based on the difference between sensory input and internal model output because these two influences worked in opposite directions to each other.

In sum, the present study found that knowledge of typical motions of objects influenced both FLE and RM, although such knowledge differently affected them. This is the first report that cognitive factors influence FLE. Many other factors can be examined and compared in both FLE and RM: eye movements (in FLE, Nijhawan 2001; in RM, Kerzel 2000; Nagai & Saiki 2006) and human internal models of physical laws (only investigated so far in RM: gravity, implied friction, see Hubbard 1995b for review). We suggest that comparing the influence of various cognitive factors on FLE and RM will lead to a better understanding of spatial errors and motion perception mechanisms.

References

Anstis, S. (2007). The flash-lag effect during illusory chopstick rotation. *Perception* **36**: 1043–1048.

Baldo, M. V. C., Kihara, A. H., Namba, J., & Klein, S. A. (2002). Evidence for an attentional component of the perceptual misalignment between moving and flashing stimuli. *Perception* **31**: 17–30.

Baldo, M. V., & Klein, S. A. (1995). Extrapolation or attention shift? *Nature* **378**: 565–566.

Brenner, E., & Smeets, J. B. (2000). Motion extrapolation is not responsible for the flash-lag effect. *Vision Res* **40**: 1645–1648.

Eagleman, D. M., & Sejnowski, T. J. (2000a). Motion integration and postdiction in visual awareness. *Science* **287**: 2036–2038.

Eagleman, D. M., & Sejnowski, T. J. (2000b). The position of moving objects. Response. *Science* **289**: 1107a.

Finke, R. A., & Freyd, J. J. (1985). Transformations of visual memory induced by implied motions of pattern elements. *J Exp Psychol Learn Mem Cogn* **11**: 780–794.

Finke, R. A., Freyd, J. J., & Shyi, G. C. W. (1986). Implied velocity and acceleration induce transformations of visual memory. *J Exp Psychol Gen* **115**: 175–188.

Finke, R. A., & Shyi, G. C. W. (1988). Mental extrapolation and representational momentum for complex implied motion. *J Exp Psychol Learn Mem Cogn* **14**: 112–120.

Freyd, J. J. (1987). Dynamic mental representations. *Psychol Rev* **94**: 427–438.

Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *J Exp Psychol Learn Mem Cogn* **10**: 126–132.

Freyd, J. J., & Finke, R. A. (1985). A velocity effect for representational momentum. *Bulletin of the Psychonomic Society* **23**: 443–446.

Freyd, J. J., & Johnson, J. Q. (1987). Probing the time course of representational momentum. *Exp Psychol Learn Mem Cogn* **13**: 259–268.

Freyd, J. J., & Miller, G. F. (1992, November). *Creature Motion.* Paper presented at the 33rd Annual Meeting of the Psychonomic Society, St. Louis, Mo.

Fröhlich, F. W. (1923). Über die Messung der Empfindungszeit [Measuring the time of sensation]. *Zeitschrift für Sinnesphysiologie* **54**: 58–78.

Halpern, A. R., & Kelly, M. H. (1993). Memory biases in left versus right implied motion. *J Exp Psychol Learn Mem Cogn* **19**: 471–484.

Hayes, A. E., & Freyd, J. J. (2002). Representational momentum when attention is divided. *Vis Cogn* **9**: 8–27.

Hubbard, T. L. (1995a). Cognitive representation of motion: Evidence for representational friction and gravity analogues. *J Exp Psychol Learn Mem Cogn* **21**: 241–254.

Hubbard, T. L. (1995b). Environmental invariants in the representation of motion: implied dynamics and representational momentum, gravity, friction, and centripetal force. *Psychonom Bull Rev* **2**: 322–338.

Hubbard, T. L. (1997). Target size and displacement along the axis of implied gravitational attraction: effects of implied weight and evidence of representational gravity. *J Exp Psychol Learn Mem Cogn* **23**: 1484–1493.

Hubbard, T. L., & Bharucha, J. J. (1988). Judged displacement in apparent vertical and horizontal motion. *Perception & Psychophysics* **44**: 211–221.

Hubbard, T. L., & Ruppel, S. E. (1999). Representational momentum and the landmark attraction effect. *Can J Exp Psychol* **53**: 242–255.

Kerzel, D. (2000). Eye movements and visible persistence explain the mislocalization of the final position of a moving target. *Vision Res* **40**: 3703–3715.

Khurana, B., & Nijhawan, R. (1995). Extrapolation or attentional shift? Reply. *Nature* **378**: 566.

Khurana, B., Watanabe, K., & Nijhawan, R. (2000). The role of attention in motion extrapolation: are moving objects ‘corrected’ or flashed objects attentionally delayed? *Perception* **29**: 675–692.

Kirschfeld, K., & Kammer, T. (1999). The Fröhlich effect: a consequence of the interaction of visual focal attention and metacontrast. *Vision Res* **39**: 3702–3709.

Mackay, D. M. (1958). Perceptual stability of a stroboscopically lit visual field containing self-luminous objects. *Nature* **181**: 507–508.

Metzger, W. (1932). Versuch einer gemeinsamen Theorie der Phänomene Fröhlichs und Hazelhoffs und Kritik ihrer Verfahren zur Messung der Empfindungszeit [An attempt toward a common theory of the phenomena of Fröhlich and Hazelhoff and a criticism of their methods to measure sensation time]. *Psychologische Forschung* **16**: 185–218.

Nagai, M., Kazai, K., & Yagi, A. (2000). Larger flash lag effect when flashed objects are presented at the onset of a moving object. *Perception* **29**: S93.

Nagai, M., Kazai, K., & Yagi, A. (2002). Larger forward displacement in the direction of gravity. *Vis Cogn* **9**: 28–40.

Nagai, M., & Saiki, J. (2005). Illusory motion and representational momentum. *Perception & Psychophysics* **67**: 855–866.

Nagai, M., & Saiki, J. (2006). Re-examination of eye-movement related factors of representational momentum. *The Japanese Journal of Psychology* **77**: 105–114 (in Japanese).

Nagai, M., & Yagi, A. (2001). Pointedness effect on representational momentum. *Memory & Cognition* **29**: 91–99.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature* **370**: 256–257.

Nijhawan, R. (1997). Visual decomposition of colour through motion extrapolation. *Nature* **386**: 66–69.

Nijhawan, R. (2001). The flash-lag phenomenon: object motion and eye movements. *Perception* **30**: 263–282.

Purushothaman, G., Patel, S. S., Bedell, H. E., & Öğmen, H. (1998). Moving ahead through differential visual latency. *Nature* **396**: 424.

Reed, C. L., & Vinson, N. G. (1996). Conceptual effects on representational momentum. *J Exp Psychol Hum Percept Perform* **22**: 839–850.

Verfaillie, K., & d’Ydewalle, G. (1991). Representational momentum and event course anticipation in the perception of implied periodical motions. *J Exp Psychol Learn Mem Cogn* **17**: 302–313.

Vinson, N. G., & Reed, C. L. (2002). Sources of object-specific effects in representational momentum. *Vis Cogn* **9**: 41–65.

Watanabe, K. (2004). Visual grouping by motion precedes the relative localization between moving and flashed stimuli. *J Exp Psychol Hum Percept Perform* **30**: 504–512.

Watanabe, K., Nijhawan, R., Khurana, B., & Shimojo, S. (2001). Perceptual organization of moving stimuli modulates the flash-lag effect. *J Exp Psychol Hum Percept Perform* **27**: 879–894.

Watanabe, K., Sato, T. R., & Shimojo, S. (2003). Perceived shifts of flashed stimuli by visible and invisible object motion. *Perception* **32**: 545–559.

22

Perceptual asynchronies and the dual-channel differential latency hypothesis

HULUSI KAFALIGÖNÜL, SAUMIL S. PATEL, HALUK ÖĞMEN,
HAROLD E. BEDELL, AND GOPATHY PURUSHOTHAMAN

Summary

The dual-channel differential latency hypothesis (Öğmen et al. 2004) successfully accounts for many aspects of the flash-lag effect (FLE). Here we use the dual-channel differential latency hypothesis to explain an illusion of perceived line length that can be viewed as one component of an illusion reported by Cai and Schlag (2001a). In the phenomenon studied here, a flash is presented collinear with a moving line that is simultaneously changing in length. The moving line is perceived to be misaligned with the flash (the FLE) and the length of the moving line is perceived to differ from its physical length at the instant of the flash. We designate this phenomenon the Cai line-Length Effect (CLE). Our analysis treats a horizontally moving line that also changes its vertical length as composed of two simultaneous motion components: (1) horizontal motion, and (2) vertical expansion or contraction. We measured perceived position misalignment and length misperception in the CLE paradigm, as well as separately for stimuli with the individual motion components of the CLE, as a function of target luminance. Perceived position misalignment and length misperception varied similarly with target luminance, both in the CLE paradigm and when the individual motion components were tested separately. The misperception of stimulus position and length in the CLE reflects an additional processing delay that may be caused by an interaction between the motion components in two directions. We conclude that the dual-channel differential latency hypothesis can account for the perceived spatiotemporal misalignment of stimulus features defined by motion components in terms of the neural latencies involved in processing these motion components.

A moving object is perceived to be displaced spatially in the direction of motion with respect to the perceived position of a flashed object, even when the two objects are aligned physically on the retina. This illusion is usually called the flash-lag effect: FLE (for reviews see Krekelberg & Lappe 2001; Nijhawan 2002; Öğmen et al. 2004). However, because the perceived displacement can change from a lag to a lead when the detectability of the flashed and the moving objects is manipulated (Purushothaman et al. 1998; Patel et al. 2000; Öğmen et al. 2004), we suggest that the flash-misalignment effect (FME) is a more appropriate designation. This illusory phenomenon is important to understand because it bears upon the temporal fidelity of sensory processing as well as on the accuracy of visually guided behaviors, such as catching a moving object (Nijhawan 1994).

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

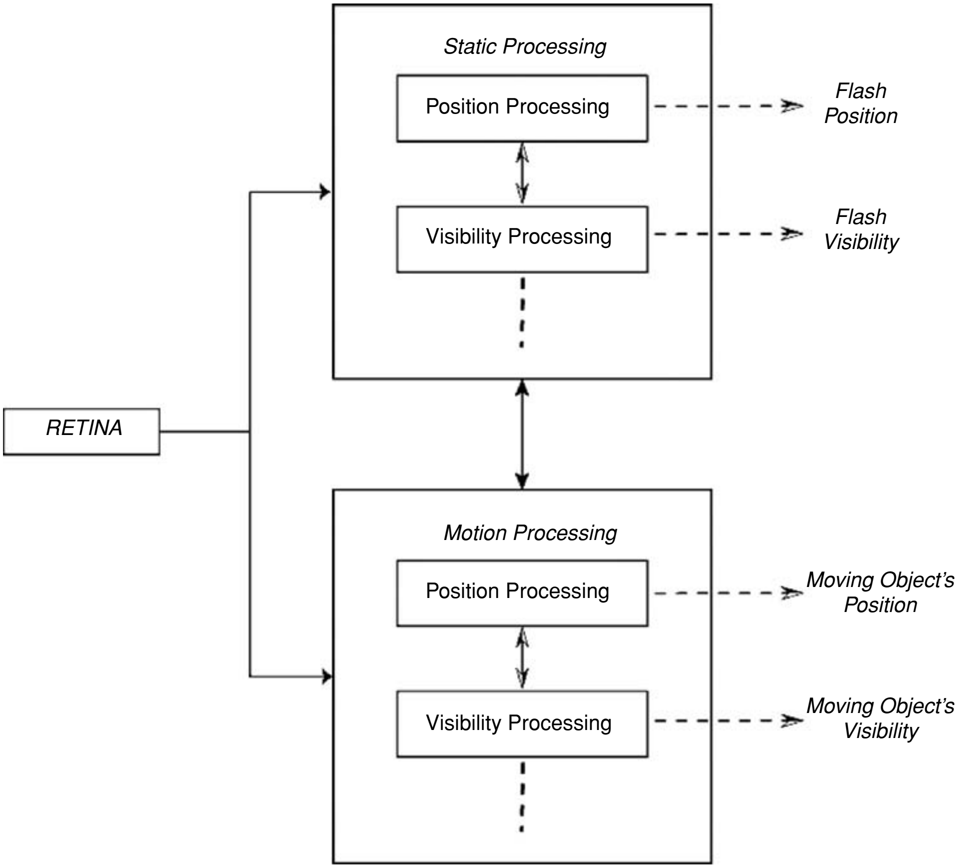


Fig. 22.1 The architecture of the dual-channel differential latency model. Static and moving stimuli are processed in different subsystems. Within each subsystem (or channel) various stimulus attributes (e.g., visibility and position) are processed by largely separate modules (adapted from Öğmen et al. 2004).

Öğmen et al. (2004) reviewed several proposed explanations for the FME. Here, we will focus primarily on the *dual-channel differential latency hypothesis,* which, in our opinion, provides the most parsimonious explanation of the FME. In this chapter, we seek to extend the dual-channel differential latency hypothesis to account for the perceptual outcome in stimulus conditions that include simultaneous motion components in two directions.

We begin by providing a brief description of the dual-channel differential latency hypothesis. The reader is referred to Öğmen et al. (2004) for a more comprehensive description. In our model, static and moving stimuli are processed in parallel by largely separate subsystems (i.e., dual channels), termed the static and motion systems, respectively. Moreover, within each system (or channel), the computations of stimulus position and stimulus visibility are different processes with different temporal properties (Fig. 22.1). These subsystems also interact to ensure that a stimulus generally produces a coherent visual percept. The time that signals from the retina take to reach the two subsystems may be different, in part because they could reside in separate anatomical areas in the brain. In addition, because the subsystems that process static and moving stimuli are likely to have different temporal properties, the resulting percepts may be delayed further by differing amounts. Similar concepts also have been applied to motion/color asynchrony phenomenon (Moutoussis & Zeki 1997; Bedell et al. 2003; Arnold 2005; Bedell et al. 2006).

In the present model, a flash-lag effect occurs if the perception of the flashed object is delayed relative to that of the moving object. The delay between the retinal stimulation and the percept that it generates consists of two components: (1) the accumulated propagation delays (T_p), and (2) the additional delays introduced by the subsystem's processing dynamics (T_d). Whereas T_p is largely independent of the temporal characteristics of the stimulus, T_d is not. In other words, if we consider a stimulus in terms of its temporal frequency components, the effect of T_p is to introduce a phase shift in each component directly proportional to the component's temporal frequency. On the other hand, T_d reflects the *interaction* between the stimulus and the processing subsystem. If the subsystem is linear shift invariant, then processing can be viewed as the convolution of the stimulus with the system's temporal impulse response function. For example, if the processing subsystem is a linear, shift-invariant, first-order, low-pass filter characterized by a single time constant (τ), and the input to this system is a moving stimulus that can be described as a linear change in position over time (i.e., a ramp stimulus), then the steady state output of the system is a ramp that is delayed by exactly the time constant, τ. On the other hand, if the input is a sinusoidal stimulus of temporal frequency f, the steady state output of the same processing subsystem will have a phase shift of $a\tan(2\pi f \tau)$ (the phase response of the system) corresponding to a delay of $a\tan(2\pi f \tau)/2\pi f$ that depends on f and τ. For a sinusoidal stimulus, the delay will be relatively constant for low temporal frequencies but will decrease and approach zero for high temporal frequencies. The model's explanation for the FME that occurs with a continuously moving object is illustrated in Fig. 22.2.

In the example shown in Fig. 22.2, the moving object is assumed to have begun its motion sufficiently early so that the *motion processing system* (or channel) and hence the trajectory of the moving object's perceived position has already reached steady state. In steady state, the total time, including the time needed to relay the retinal information to the position computation subsystem in the motion processing system and the time that the subsystem takes to compute the position, equals d_m. Hence, the *perceived* position of the moving object is delayed by d_m. Note that when the motion processing system is in steady state, the visibility computation subsystem does not interact with the position computation subsystem and hence is largely ignored in the present analysis for the perceived position of the moving object. The stationary flash is presented when it is physically collinear with the position of the moving object. In the dual-channel differential latency model, the stationary flash's position and visibility are computed by a separate *static processing system* (or channel). First, we examine the visibility processing subsystem in the static processing system. The total time to relay the flash's information from the retina to the visibility computation subsystem in the static processing system and the time that the subsystem

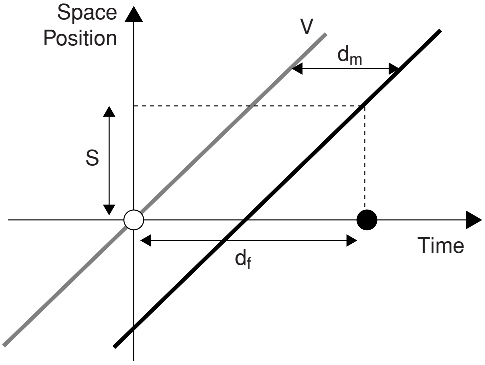


Fig. 22.2 Space–time diagram that explains the FME based on the dual-channel differential latency model. In this example, the moving object is assumed to have begun its motion sufficiently early such that the trajectory of its perceived position has reached steady state. The *physical* position of the moving object is represented by the oblique gray line. The *perceived* position of the moving object, which is delayed by d_m, is represented by the black oblique line. The flash (unfilled circle) is presented when it is physically collinear with the position of the moving object. If the flash is perceived after a delay of d_f (filled circle), the perceived position of the moving line is ahead of the perceived position of the flash by a spatial displacement, s. The value of s is equal to $v(d_f - d_m)$, where v is the speed of the moving object and $(d_f - d_m)$ is the differential latency between the perception of the flashed and the moving object. Note that even though the visibility processing of the stationary flash is dynamic, the filled circle in the figure only shows the instant when the output crosses a critical signal-to-noise ratio for the first time.

takes to produce an output larger than a critical signal-to-noise ratio is equal to d_f. Hence, the flash is perceived after a delay of d_f. The position computation subsystem in the static processing system also operates in parallel to yield the perceived position of the stationary flashed object. For the analysis here, we assume that a steady state output of the position computation subsystem is available before the output of the visibility processing subsystem reaches the critical signal-to-noise ratio. In the example in Fig. 22.2, because d_f is greater than d_m, the perceived position of the moving line is ahead of the perceived position of the flash by a spatial displacement, s. The value of s is equal to $v(d_f - d_m)$, where v is the speed of the moving object and $(d_f - d_m)$ is the differential latency between the perception of the flashed and the moving object. Note that in the dual-channel differential latency model the space–time characteristics of the perception of a stationary flashed and a moving object can be represented independently in space–time plots such as the one shown in Fig. 22.2. The reader is referred to Öğmen et al. (2004) for a comprehensive review of the transient dynamics of the motion processing system.

It is important to recognize that Fig. 22.2 is ambiguous in that one can view the perceived trajectory of the moving object as either temporally or spatially lagging the physical stimulus. To resolve this ambiguity, we examined the perceived position of the moving object from the onset of motion to the time when the perception of position reaches steady state (Öğmen et al. 2004). The results indicate that the perceived position of the moving object lags the physical retinal stimulation *temporally.* This conclusion is consistent with the outcome of other studies that introduced changes in the speed (Brenner & Smeets 2000) or direction of the moving object's trajectory (Eagleman & Sejnowski 2000; Whitney et al. 2000).

Most of the studies that investigated FME used one stimulus that moved in a single linear or circumferential direction and another stimulus that was flashed. However, the dual-channel differential latency hypothesis can be extended to predict the outcome of experiments with stimuli that contain more complex motion components. Cai and colleagues (Cai et al. 2000; Cai & Schlag 2001a; Cai & Schlag 2001b; Cai & Cavanagh 2002; Cai 2003) introduced an experimental paradigm in which a moving target changed features such as length and color either smoothly or abruptly. In one of these experiments, the stimulus consisted of a moving line that gradually changed its length as it moved. In the middle of the motion trajectory, a brief flash was presented collinear with the moving line. As expected on the basis of the FME, the observer sees the moving line spatially ahead of the flash's position at the instant (s)he perceives the flash. In addition, the observer perceives the length of the moving line to be longer than its physical length at the time the flash is presented. We will call this illusion the Cai line-Length Effect (CLE). Our analysis treats a horizontally moving line that also changes its vertical length as a combination of two simultaneous motion components: (1) horizontal motion, and (2) vertical expansion or contraction. It is important to note the distinction between a stimulus resulting in the CLE and a stimulus with circular motion. In both stimuli, assuming stimulus motion occurs in a frontoparallel plane, there are two linear (horizontal and vertical) components of motion. However, in the former case, the components of motion can be perceived separately in the frontoparallel plane and/or can be integrated into a single direction of depth toward or away from the frontoparallel plane, whereas in the latter case they are integrated into a single direction in the frontoparallel plane. In other words, the processing of the stimulus resulting in the CLE and a stimulus in circular motion is distinct and hence can take different amounts of time. We will visit this issue again toward the end of the chapter. Here, we consider how the CLE can be explained by the dual-channel differential latency hypothesis.

Consider a vertical line that starts moving rightward from a location in the left visual field at a speed of v_h deg/sec. As the line moves its length increases at a rate of v_v deg/sec. Both the rightward motion and the increase in line length stop when the line reaches a location in the right visual field. The observer fixates a position that is horizontally halfway between the starting and the ending positions of the line's rightward motion trajectory and vertically just below the moving line. When the moving line is directly above the fixation point, a flash is presented just below the fixation point in physical horizontal alignment with the moving line. The predictions of the dual-channel differential latency hypothesis for this version of the CLE paradigm are illustrated in Fig. 22.3. These predictions are derived for one level of detectability of the moving line but can be generalized readily to other levels of detectability (see Öğmen et al. 2004 for modifications relating to detectability).

The time-varying aspects of the stimulus are decomposed into two components: (1) a component that corresponds to horizontal motion, and (2) a component that corresponds to vertical expansion (i.e., an increase in length). For simplicity, we assume for now that the rightward line motion and the vertical expanding motion are processed separately without interactions. Figure 22.3 only shows the steady state responses corresponding to these two

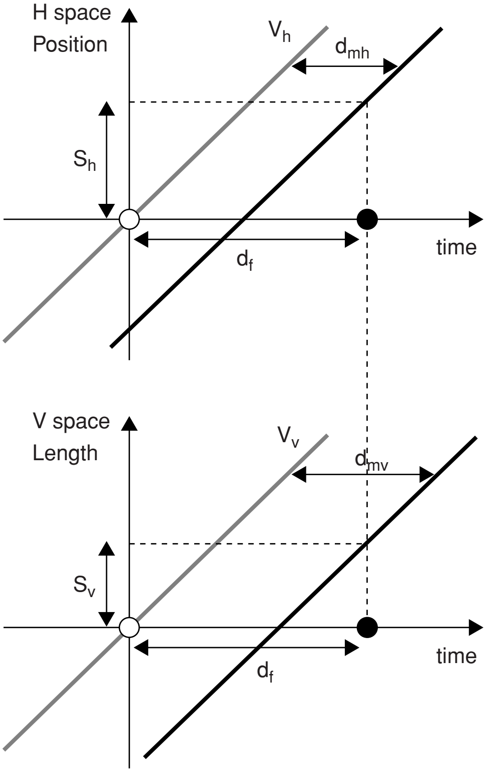


Fig. 22.3 Space–time diagram illustrating the CLE. The top and bottom parts of this figure, which represent the simultaneous horizontal position and vertical length of the stimulus, respectively, are similar to Fig. 22.2. Note that for the bottom figure, the origin's y-coordinate represents a nonzero length. For purposes of generalizability, the steady state latencies for the horizontal position and vertical length perception are d_{mh} and d_{mv} respectively. When the flash is perceived (filled circle), the perceived position of the moving object is s_h ahead of the flash and the perceived length of the moving object is s_v longer than the physical line length at the instant of flash presentation. The values, s_h and s_v, are equal to $v_h(d_f - d_{mh})$ and $v_v(d_f - d_{mv})$, respectively, where v_h is the horizontal speed of the moving line and v_h is its rate of length change. In this illustration and in our experiments, v_h and v_v are equal. The differential latencies between the perception of the flashed and the moving objects' horizontal position and vertical length are $(d_f - d_{mh})$ and $(d_f - d_{mv})$, respectively.

types of stimulus motion. As seen in Fig. 22.3, each type of motion may be delayed by a different amount depending on the dynamics of the relevant processing mechanism. The following predictions can be derived for the perception of the moving stimulus at the instant that the flash becomes visible:

(1) When the flash is perceived, the horizontal position of the moving line will be misaligned, as expected from the FME. The sign and magnitude of the misalignment will depend on the relative delays involved in processing the horizontal position of the moving line and the detection of the flash.

(2) When the flash is perceived, the length of the moving line will be misperceived compared to its physical length at the time of flash presentation, as described by the CLE. The sign and magnitude of the misperception will depend on the relative delays involved in processing the vertical length of the changing line and the detection of the flash.
(3) Increasing the detectability of the moving stimulus should reduce its latency (Roufs 1974; Williams & Lit 1983) and should therefore cause the perceived position misalignment in the horizontal direction (the magnitude of the FME) to change in the direction of a greater position lead for the moving stimulus.
(4) Increasing the detectability of the moving stimulus should also cause the misperception of line length to change in the direction of a greater length lead.
(5) If the two types of motion are processed largely independently and the perceptual illusion related to each type of motion occurs due to differential processing latencies, then the effects of changing stimulus detectability in the CLE paradigm should be equivalent to the effects observed when the horizontal and vertical (length-change) components of line motion are examined separately.

We tested these predictions in a series of psychophysical experiments. Because the temporal response of the visual system depends on retinal eccentricity (McKee & Taylor 1984; Tyler 1985), in all of these experiments the retinal eccentricity, at which the critical information for the observers' judgments occurred, was kept constant.

In our first experiment, we measured the spatial misalignment between a flashed bar and a line of fixed vertical length that moved in the horizontal direction. This experiment examined the horizontal motion component of the CLE. The stimulus configuration for this experiment is shown in Fig. 22.4.

In separate runs, we measured the flash misalignment for a moving line that was on average 2.1 and 3.2 log units (LU, averaged across three observers, one of whom was naive) above its detection threshold. The results of this experiment are shown in Fig. 22.5.

For both levels of detectability, the moving line was perceived to be ahead of the flash at the instant the flash was perceived, that is, a positive FME or a flash-lag effect was observed. Consistent with previous results (Purushothaman et al. 1998; Patel et al. 2000; Öğmen et al. 2004), the perceived spatial position lead of the moving object increases with an increase in its detectability ($F[1,6] = 38.7, p = 0.007$). The rate of increase of position is about 30 msec per LU of the moving line's detectability. We will use this rate in the following to compare the steady state dynamics of position processing for targets that move in the horizontal direction and that expand/contract in the vertical direction.

A repeated-measures analysis of variance (ANOVA) was performed to statistically evaluate the results of this and subsequent experiments. The factors included in the ANOVA were luminance (two levels) and experiment (four levels). The outcome variable was the magnitude of the FME, in min arc. The main effects of luminance ($F[1,2] = 85.9, p = 0.01$) and experiment ($F[3,6] = 18.3, p = 0.005$) were significant, but their interaction was not ($F[3,6] = 0.6, p = 0.65$). Posthoc contrasts were used to compare specific pairs of conditions within and across the separate experiments (see the following). A second repeated-measures ANOVA using the same factors was performed to evaluate possible changes in the inverse slopes of the psychometric functions (outcome variable), specified

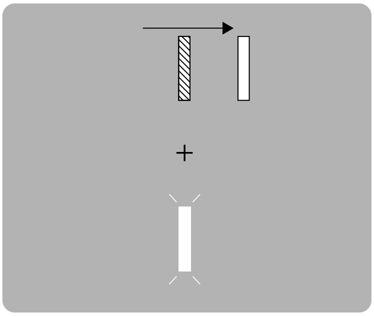


Fig. 22.4 Stimulus configuration for the first experiment. A vertical line (8.8 × 88 min arc) moved horizontally along a trajectory that was 1 deg (from the center of the fixation target to the lower edge of the line) above a fixation cross (1.1 cd/sq-m) on a computer monitor (pixel size = 2.2 min arc) connected to a VSG-2/3 board. The background luminance of the screen was 1.9 cd/sq-m. The speed of the moving line was 1.9 deg/sec and the duration of motion was 1.3 sec (69 frames of 18.75 msec each). The direction of motion was chosen randomly for each trial. When the position of the moving line was in the vicinity of the fixation cross (textured rectangle above fixation cross), a vertical bar (white 8.8 × 88 min arc rectangle below fixation cross) was flashed 1 deg (from the center of the fixation cross to the upper edge of the bar) below it. The duration of the flash was 18.75 msec. The luminance of the flash was 10.2 cd/sq-m and its detectability averaged across the three observers was 1.45 LU (± 0.30 SD) above its detection threshold. The viewing distance was 1 m. The time instant at which the flash was presented varied from trial to trial using the method of constant stimuli. The observers binocularly judged whether the position of the moving line was to the left or right of the flashed bar. A psychometric function was constructed from the collected data, and the 50% point on the curve corresponds to the point of *subjective* alignment (PSA). The *physical* misalignment occurring at the PSA was defined as the flash misalignment or, equivalently, as the misalignment in the position of the moving line. In the FME, at the time the observers perceive the flash, the moving line is perceived to be ahead of its physical location, as illustrated by the white rectangle above the fixation cross.

in min arc per 1 SD change in the percentage of responses. There was no effect of luminance ($F[1,2] = 0.002$, $p = 0.97$) or experiment ($F[3,6] = 6.3$, $p = 0.09$), and no evidence of a luminance by experiment interaction ($F[3,6] = 0.09$, $p = 0.84$).

In our second experiment, we examined the vertical component of motion in the CLE, that is, the vertical change in line length. The stimuli for this experiment are shown in Fig. 22.6.

The perceived length at the time of the flash was measured for two levels of detectability of the expanding and contracting lines. To obtain a bias-free estimate, the data for the expanding- and contracting-line conditions were averaged. The average data for the same observers are shown in Fig. 22.7.

As in the first experiment, the perceived length of the line is *ahead* of its physical length at the instant the flash is perceived for both detectabilities of the changing line. These results indicate that an FME also occurs for the line-length component of the CLE.

In agreement with Experiment 1, the lead in the length of the changing line increases as the level of detectability increases ($F[1,6] = 30.9, p = 0.009$). The rate of increase in length

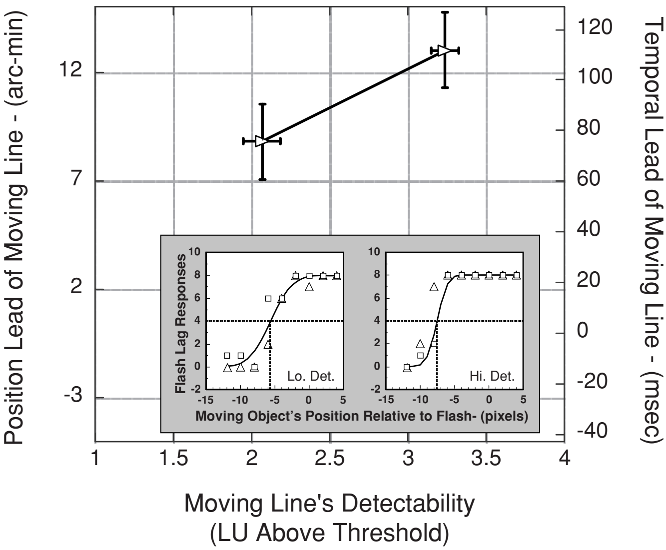


Fig. 22.5 The misalignment in the position of the moving line as a function of its detectability in Experiment 1. The left (right) ordinate represents the position misalignment in spatial (temporal) units. The temporal misalignment is obtained by dividing the spatial misalignment by the speed of the moving line. The luminance of the moving lines was 4.9 and 62.0 cd/sq-m in the low and high detectability conditions, respectively. The plotted data are also equivalent to the FME with positive numbers representing a flash-lag. The data shown are the average of three observers (one of whom was naive). In this and other figures, error bars represent the standard deviation across observers. ***Inset:*** The psychometric functions of the naive observer for two detectability conditions. The x-axis represents the position of the moving line relative to the position of the flash at the time of the presentation of the flash (-ve numbers represent a spatial lag for the moving line). The y-axis represents the number of times an observer responded that the flash spatially lagged the moving line. Each relative position is represented eight times in an experimental run. In this and subsequent figures, the data shown (squares and triangles) are from two experimental runs and the fitted curves represent cumulative Gaussian functions. The vertical dashed lines represent the position of the moving line at the time of the flash presentation that resulted in a perceived spatial alignment of the flash and the moving line. The inverse slopes (±SEM) of the psychometric functions averaged across observers for low and high detectability conditions are 2.30 ± 0.34 and 2.48 ± 0.68 pixels respectively. Note that in this and all other data plots, the sizes of the symbols are different only to make them legible when overlapping with each other, and slope values are given in pixel units.

lead is about 26 msec per LU change in detectability, which is similar to that obtained in Experiment 1 for the horizontal component of the CLE. However, in comparison to the results of Experiment 1 shown in Fig. 22.5, the data in Fig. 22.7 are shifted downward by approximately 30 msec ($F[1,6] = 5.7$, $p = 0.1$). This shift can be interpreted as a longer delay for processing the vertical expansion/contraction compared to the horizontal position of the changing or moving line. As indicated by the example psychometric functions in the inset to Fig. 22.7, there is also a substantial asymmetry in the FME for the line length of expanding and contracting lines. The direction of this asymmetry is consistent with that

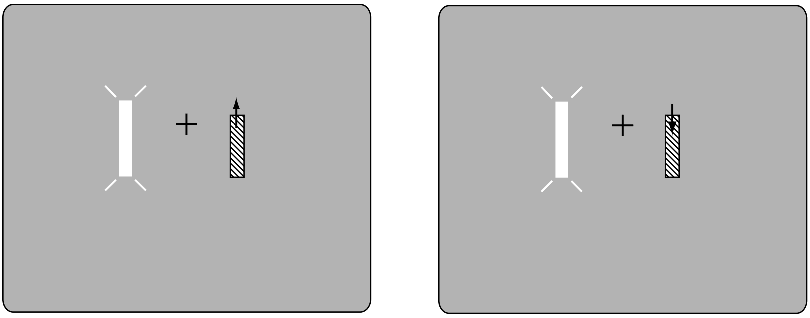


Fig. 22.6 Stimulus conditions for Experiment 2. A *stationary* line (textured rectangle) that either expanded (left panel) or contracted smoothly in length (right panel) was presented 1 deg to the right (from the center of the fixation cross to the horizontal center of the line) of a fixation cross. The bottom of the changing line remained at a fixed position, 1.6 deg below the center of the fixation cross. When the upper edge of the expanding or contracting line was aligned with the center of the fixation cross, a bar (white rectangle) was flashed 1 deg to the left of the fixation cross. The length of the line when the flash occurred was 88 min arc. From trial to trial the length of the *flashed* bar was varied according to the method of constant stimuli. Because detectability of a flashed object affects the FME (Purushothaman et al. 1998; Patel et al. 2000; Öğmen et al. 2004), it is important to note that for the range of flash bar lengths used in this experiment, the detectability of the flashed bar changed negligibly (~0.1 LU). The expanding and contracting line-length conditions were run in separate blocks of trials. The observer's task was to judge whether the flashed bar was longer or shorter than the line that changed smoothly in length. The perceived length of the line at the time that the flash was perceived was defined as the 50% point on the psychometric function.

found in a recent study of the FLE in depth (Harris et al. 2006). However, in spite of the change in line length that occurred in Experiment 2, the observers did not report a strong perception of motion in depth.

In the third experiment, we measured the perceived misalignment of the horizontal position of the line in the CLE paradigm. If the processing of horizontal motion and vertical length change do not interact, then the perceived horizontal position misalignment of a line that simultaneously moves horizontally and changes length vertically (i.e., a *changing line*) should be identical to the perceived position misalignment for the moving line obtained in Experiment 1. In other words, the FME should not be affected by what happens to the length of the line. The stimulus used in this experiment is shown in Fig. 22.8, and the data are presented in Fig. 22.9.

Consistent with the results of Experiment 1, the line is perceived to be ahead of its physical position at the instant the flash is perceived, that is, an FME is observed for the horizontal position of the line that simultaneously moves and changes length in the CLE paradigm. The position-lead of the changing line increases with an increase in its detectability ($F[1,6] = 36.1$, $p = 0.007$) at a rate of about 25 msec per LU of detectability. This rate of change of position-lead is similar to that observed in the first experiment. However, compared to the results of the first experiment shown in Fig. 22.2, the data in Fig. 22.9 are shifted down significantly, by an amount equal to approximately 60 msec ($F(1,6) = 24.4$,

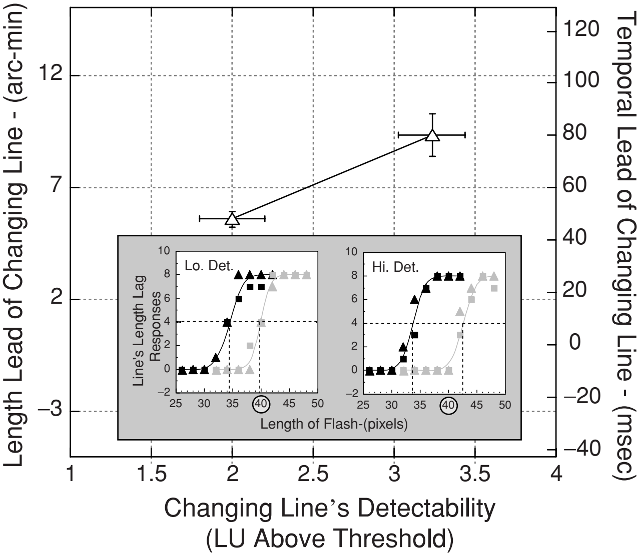


Fig. 22.7 The misperception in length of a line as function of the line's detectability in Experiment 2. For an expanding (contracting) stimulus, the vertical axis represents the amount by which the perceived length of the line is longer (shorter) than the line length that was presented at the time the flash was presented. The data shown are the average of three observers (one of whom was naive). ***Inset:*** The psychometric functions of the naive observer for two detectability conditions. The x-axis represents the length of the flash. The physical length of the line at the time of the flash presentation was forty pixels (88 min arc). The y-axis represents the number of times an observer responded that the length of the line was shorter than that of the flash. Each flash's length is represented eight times in an experimental run. The black (gray) symbols and curve correspond to the expanding (contracting) line condition. The vertical dashed lines represent the length of the flash that resulted in a perceived length match of the flash and the expanding/contracting line. The inverse slope (±SEM) of the psychometric functions averaged across observers for low and high detectability conditions of expanding (contracting) lines are 1.69 ± 0.23 (1.62 ± 0.15) and 1.76 ± 0.05 (1.43 ± 0.13) pixels respectively.

$p = 0.03$). We will discuss this 60 msec shift after we report the results of the fourth experiment.

Finally, in our fourth experiment, we measured the CLE using the stimulus shown in Fig. 22.10. The data are shown in Fig. 22.11.

For the lower detectability of the line, its perceived length *lags* the physical length at the instant the flash is perceived. In other words, the misperception of line length in this condition is consistent with a flash-*lead* for the vertical component of motion in the CLE. As mentioned above, a flash-lead is encountered also in experiments similar to Experiment 1 when the delay in processing the flash becomes smaller than the delay to process the moving line (Purushothaman et al. 1998; Patel et al. 2000; Öğmen et al. 2004). Consistent with the data from our previous three experiments, the perceived length of the line shifts in the direction of a relative *lead* as the detectability of the changing line increases

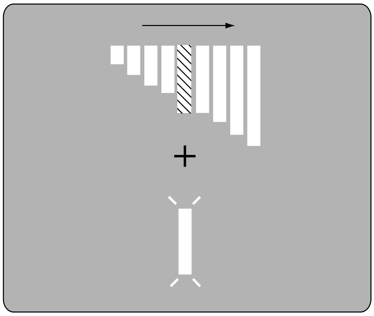


Fig. 22.8 The stimulus configuration for Experiment 3. A line moved horizontally and also expanded in length smoothly above the fixation cross. The direction of motion was chosen randomly for each trial. When the moving/expanding line was in the vicinity of the fixation cross (textured rectangle), a bar was flashed below the fixation cross. The top of the changing line remained at a fixed position, 2.6 deg above the center of the fixation cross. The bottom of the changing line was approximately 1 deg above the fixation cross as in Experiment 1. The dimensions and eccentricity of the flashed bar were identical to those in Experiment 1. The horizontal speed of the moving line was identical to that in Experiment 1, and its rate of expansion was identical to that in Experiment 2. The experimental procedure, observer's judgment, and data analysis were identical to those in Experiment 1. We only used the expanding-line condition for this experiment.

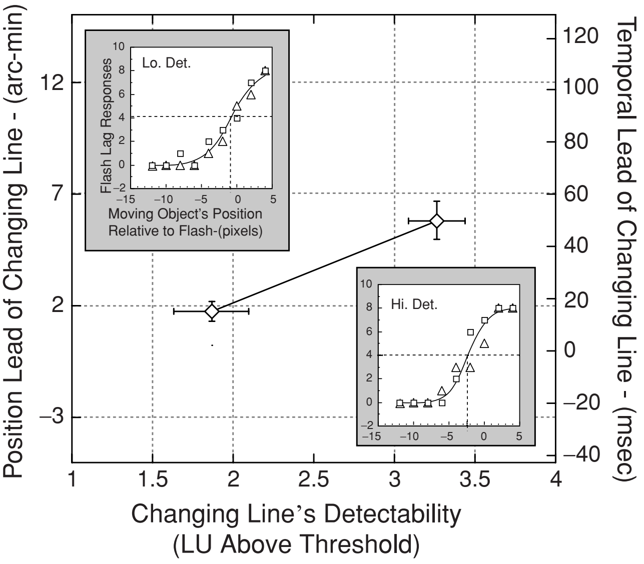


Fig. 22.9 The perceived misalignment in the position of the changing line as a function of its detectability in the CLE paradigm, in Experiment 3. The inverse slope (±SEM) of the psychometric functions averaged across observers ($n = 3$) for low and high detectability conditions are 3.22 ± 0.31 and 3.27 ± 0.69 pixels respectively. ***Inset:*** The psychometric functions of the naive observer for two detectability conditions. The results are depicted using the same notation as in Experiment 1.

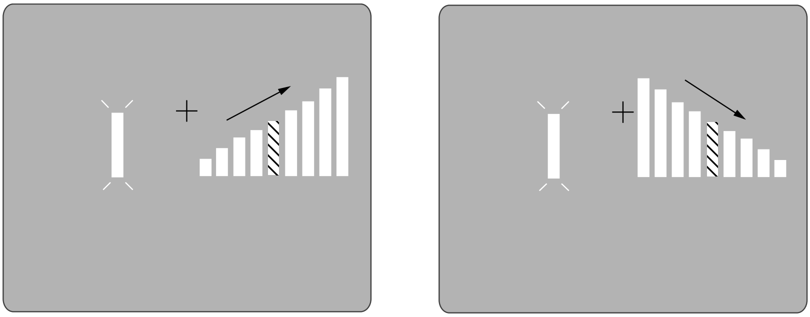


Fig. 22.10 Stimulus conditions in Experiment 4. A rightward *moving* line that either expanded (left panel) or contracted smoothly in length (right panel) was presented to the right (from the center of the fixation cross to the horizontal center of the line) of a fixation cross. The horizontal speed was identical to that in Experiments 1 and 3. The expansion/contraction rate was identical to that in Experiment 2. When the upper edge of the expanding or contracting line (textured rectangle on the right side of fixation) was approximately aligned with the center of the fixation cross, a bar was flashed 1 deg to the left of the fixation cross. The eccentricity of the flashed bar was identical to that in Experiments 1 through 3. The length of the expanding or contracting line when the flash occurred was 88 min arc. The procedure, observer's judgment, and data analysis were identical to those in Experiment 2. Expanding and contracting line conditions were run in separate blocks of trials. Because a previous study (Kanai et al. 2004) indicated that there is a significant difference between the flash-lag values of foveo–petal and foveo–fugal conditions, only one direction of motion (foveo–fugal) was used in this experiment.

($F[1,6] = 53.8$, $p = 0.004$). The rate of increase in length-lead is about 30 msec per LU change in detectability. This rate of change in length-lead as a function of the line's detectability in the CLE paradigm is similar to that in Experiment 2. However, the data in Fig. 22.11 are also shifted down relative to those in Fig. 22.7 by approximately 55 msec ($F[1,6] = 20.8$, $p = 0.03$). This downward shift in the CLE data can be attributed to an additional delay in processing, presumably because of the simultaneous presence of motion components in the vertical and horizontal directions. A comparison of the results from Experiments 3 and 4 indicates that the additional delay when more than one motion component is present is similar for the perceived length of the vertically expanding or contracting line (~55 msec) and for its perceived horizontal position (~60 msec). As indicated by the example of psychometric functions in the inset to Fig. 22.11, there is a slight asymmetry in the FME for the expanding and contracting line-length components of the CLE, but it is greatly reduced from that observed in Experiment 2.

Although the effect of changing the detectability of the stimulus is consistent with our previous results and the predictions of the dual-channel differential latency hypothesis, our data indicate a substantial decrease in FME when more than one component of motion is present in the stimulus simultaneously. This outcome indicates that our initial treatment of vertical expansion/contraction and horizontal motion processing as noninteractive was an oversimplification. Within the framework of the dual-channel differential latency hypothesis, an additional processing delay is introduced when vertical expansion/contraction and

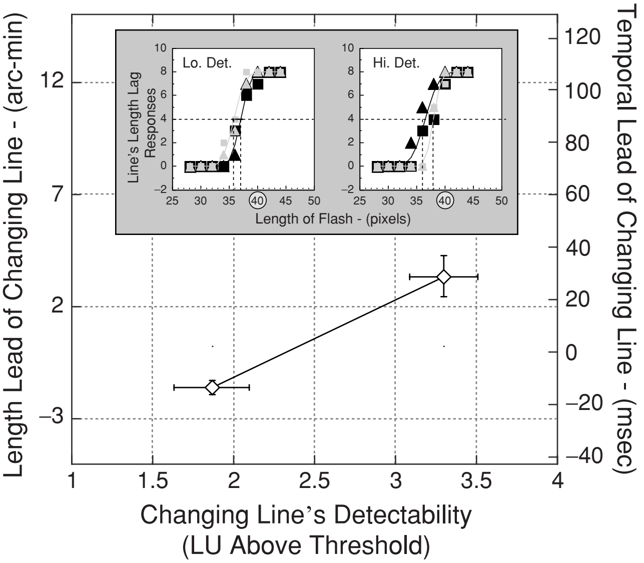


Fig. 22.11 The misperception of line length in the CLE paradigm as a function of the detectability of the line, in Experiment 4. The inverse slope (±SEM) of the psychometric functions averaged across observers ($n = 3$) for low and high detectability conditions of expanding (contracting) lines are 2.37 ± 0.68 (1.69 ± 0.17) and 2.02 ± 0.52 (1.78 ± 0.06) pixels, respectively. ***Inset:*** The psychometric functions of the naive observer for two detectability conditions. The results are depicted using the same notation as in Experiment 2.

horizontal motion components are simultaneously present (for a detailed explanation see Fig. 22.12). Moreover, the asymmetry in the magnitude of perceived length for expansion and contraction conditions (Figs. 22.7 and 22.11) implies that figural changes such as expansion/contraction coupled with motion may induce additional processes related to form–motion interactions. Previously, Regan and Beverley (1978) showed that distinct, independently adaptable motion mechanisms process looming and zooming stimuli that are perceived to move toward or away from the observer, as compared to stimuli that only translate in a frontoparallel plane. Although a detailed analysis of this additional processing delay is beyond the scope of this chapter, some speculation about mechanisms is nevertheless possible. Motion integration is a process whereby local motion signals belonging to an object are combined to determine the global motion characteristics of that object (e.g., Adelson & Movshon 1982; Grzywacz & Yuille 1991; Laurenceau & Shiffrar 1992; Mingolla et al. 1992; Wallach 1995; Pack et al. 2001). In general, local motion components of an object suffer from the "aperture problem," that is, they only can provide ambiguous information about the object's movement direction (Fennema & Thompson 1979; Marr & Ullman 1981). But line ends and terminators of the object typically signal the veridical object direction. Thus, in the stimuli shown in Figs. 22.4 and 22.6 as well as in rigidly translating or rotating lines of fixed length, the direction of the line ends unambiguously signal the direction of motion in the scene. But in the CLE stimulus (Fig. 22.8), the line end

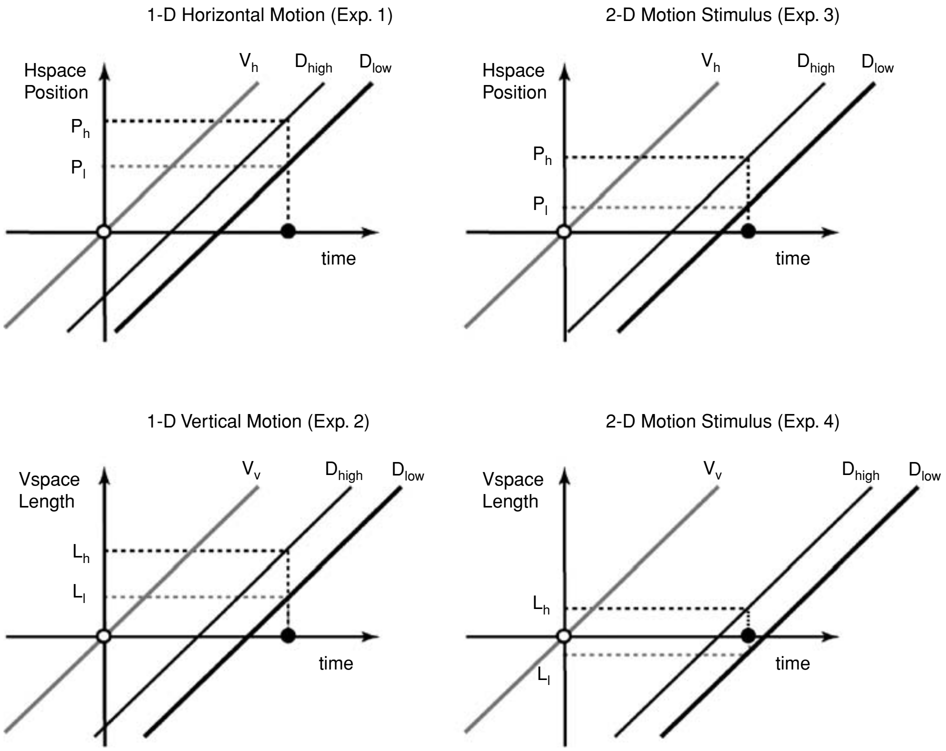


Fig. 22.12 Qualitative explanation of the results of Experiments 1–4, based on the dual-channel differential latency hypothesis. In all panels, the motion of the stimulus is represented by the oblique gray line. v_h and v_v represent the speed of the stimulus in the horizontal and vertical directions, respectively. The trajectories of perceived stimulus position are represented by the thin and thick oblique black lines, which signify low and high levels of stimulus detectability, respectively. The unfilled and filled circles represent the physical and perceived times of the flashed target, respectively. P_l and P_h represent the perceived horizontal position misalignment for low and high detectability stimuli, and L_l and L_h represent the misperception of length for low and high detectability stimuli.

signals a diagonal direction different from the horizontal direction in which the bar moves. It is possible that the additional processing delay may be the result of increased complexity in computing the global object direction from the three ambiguous local motion signals while maintaining separate perceptual identities for each component. Our data show that this additional delay is between 55 and 60 msec.

The explanation of the CLE based on the dual-channel differential latency hypothesis demonstrates that the parallel and distributed nature of neural processing may cause spatial features that are defined by motion signals to be misperceived. Although other explanations for CLE are possible, the dual-channel differential latency hypothesis provides a parsimonious explanation for the flash-lag effect (e.g., Nijhawan 1994), the Hess and Pulfrich effects (Williams & Lit 1983), motion/color asynchrony phenomena (e.g., Moutoussis &

Zeki 1997; Arnold et al. 2001; Nishida & Johnston 2002; Bedell et al. 2003; Arnold 2005; Bedell et al. 2006) as well as the CLE (Cai & Schlag 2001). The results presented here add to the overwhelming evidence for a lack of compensation of temporal asynchronies that arise in various parts of the neuronal pathways leading to perception. Consistent with previous suggestions (Purushothaman et al. 1998; Kerzel & Gegenfurtner 2003), the results presented in this chapter reinforce the need for visually guided behavior to rely less on predictive perceptual processing and more on predictive sensorimotor processing. In terms of the example of catching a moving ball, the brain is likely to compensate for the delay in the visual information by moving the arm to a predicted future spatial location.

Acknowledgments

This work was supported by R01 EY05068, R01 MH49892, and R01 EY12810.

References

Adelson, E. H., & Movshon, J. A. (1982). Phenomenal coherence of moving visual pattern. *Nature* **300**: 523–525.

Arnold, D. (2005). Perceptual pairing of colour and motion. *Vision Res* **45**: 3015–3026.

Arnold, D. H., Clifford, C. W., & Wenderoth, P. (2001). Asynchronous processing in vision: color leads motion. *Curr Biol* **11**: 596–600.

Bedell, H. E., Chung, S. T. L., Öğmen, H., & Patel, S. S. (2003). Color and motion: which is the tortoise and which is the hare? *Vision Res* **43**: 2403–2412.

Bedell, H. E., Patel, S. S., Chung, S. T. L., & Öğmen, H. (2006). Perceptual consequences of timing differences within parallel feature-processing systems in human vision. In H. Öğmen & B. G. Breitmeyer (eds.), *The First Half Second: The Microgenesis and Temporal Dynamics of Unconscious and Conscious Visual Processes* (245–258). Cambridge, MA: MIT Press.

Brenner, E., & Smeets, J. B. J. (2000). Motion extrapolation is not responsible for the flash-lag effect. *Vision Res* **40**: 1645–1648.

Cai, R. H. (2003). The Fröhlich effect is not due to a failure to perceive the beginning portion of motion trajectory *J Vis* **3**: 485a.

Cai, R. H., & Cavanagh, P. (2002). Motion interpolation of a unique feature into stimulus gaps and blind spots. *J Vis* **2**: 30a.

Cai, R. H., & Schlag, J. (2001a). A new form of illusory conjunction between color and shape. *J Vis* **1**: 127a.

Cai, R. H., & Schlag, J. (2001b). Asynchronous feature binding and the flash-lag illusion. *Invest Ophthalmol Vis Sci* **42**: S711.

Cai, R. H., Schlag-Rey, M., & Schlag, J. (2000). Displacement of the moving bar exists in the flash-lag effect. *Society for Neuroscience Abstract* **561**: 1.

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science* **287**: 2036–2038.

Fennema, C. L., & Thompson, W. B. (1979). Velocity determination in scenes containing several moving objects. *Computer Graphics and Image Processing* **9**: 301–315.

Grzywacz, N., & Yuille, A. (1991). Theories for the visual perception of local velocity and coherent motion. In J. Landy & J. Movshon (eds.), *Computational Models of Visual Processing* (231–252). Cambridge, MA: MIT Press.

Harris, L. R., Duke, P. A., & Kopinska, A. (2006). Flash lag in depth. *Vision Res* **46**: 2735–2742.

Kanai, R., Sheth, B. R., & Shimojo, S. (2004). Stopping the motion and sleuthing the flash-lag effect: spatial uncertainty is the key to perceptual mislocalization. *Vision Res* **44**: 2109–2128.

Kerzel, D., & Gegenfurtner, K. R. (2003). Neuronal processing delays are compensated in the sensorimotor branch of the visual system. *Curr Biol* **13**: 1975–1978.

Krekelberg, B., & Lappe, M. (2001). Neuronal latencies and the position of moving objects. *Trends Neurosci* **24**: 335–339.

Laurenceau, J., & Shiffrar, M. (1992). The influence of terminators on motion integration across space. *Vision Res* **32**: 263–273.

Marr, D., & Ullman, S. (1981). Directional selectivity and its use in early visual processing. *Proc R Soc Lond B Biol Sci* **211**: 151–180.

McKee, S. P., & Taylor, D. G. (1984). Discrimination of time: comparison of foveal and peripheral sensitivity. *J Opt Soc Am A* **1**: 620–627.

Mingolla, E., Todd, J. T., & Norman, J. F. (1992). The perception of globally coherent motion. *Vision Res* **32**: 1015–1031.

Moutoussis, K., & Zeki, S. A. (1997). Direct demonstration of perceptual asynchrony in vision. *Proc R Soc Lond B Biol Sci* **264**: 393–399.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature* **370**: 256–257.

Nijhawan, R. (2002). Neural delays, visual motion and the flash-lag effect. *Trends Cogn Sci* **6**: 387–393.

Öğmen, H., Patel, S. S., Bedell, H. E., & Camuz, K. (2004). Differential latencies and the dynamics of the position-computation process for moving targets, assessed with the flash-lag effect. *Vision Res* **44**: 2109–2128.

Pack, C. C., Berezovskii, V. K., & Born, R. T. (2001). Dynamic properties of neurons in cortical area MT in alert and anaesthetized macaque monkeys. *Nature* **414**: 905–908.

Patel, S. S., Öğmen, H., Bedell, H. E., & Sampath, V. (2000). Flash-lag effect: differential latency, not postdiction. *Science* **290**: 1051a.

Purushothaman, G., Patel, S. S., Bedell, H. E., & Öğmen, H. (1998). Moving ahead through differential latency. *Nature* **396**: 424.

Regan, D., & Beverley, K. I. (1978). Looming detectors in human visual pathway. *Vision Res* **18**: 415–421.

Roufs, J. A. (1974). Dynamic properties of vision – V. *Vision Res* **14**: 853–869.

Tyler, C. W. (1985). Analysis of visual modulation sensitivity. II. Peripheral retina and the role of photoreceptor dimensions. *J Opt Soc Am A* **2**: 393–398.

Wallach, H. (1995). Über Visuell Wahrgenommene Bewegungrichtung. *Psychologische Forschung* **20**: 325–380.

Whitney, D., Murakami, I., & Cavanagh, P. (2000). Illusory spatial offset of a flash relative to a moving stimulus is caused by differential latencies for moving and flashed stimuli. *Vision Res* **40**: 137–149.

Williams, J. M., & Lit, A. (1983). A luminance-dependent visual latency for the Hess effect, the Pulfrich effect, and simple reaction time. *Vision Res* **23**: 171–179.

23

Paying attention to the flash-lag effect

MARCUS V. C. BALDO AND STANLEY A. KLEIN

Summary

In the flash-lag effect (FLE) a stationary flash is usually mislocalized as lagging behind a moving object in spatiotemporal alignment. Nijhawan, who postulated a mechanism of perceptual extrapolation of motion to explain the phenomenon, rediscovered this perceptual effect. The first challenge to the motion extrapolation hypothesis included an attentional shift mechanism as the alternative, which implicitly relied on the spotlight metaphor for visual attention. Other explanations have been forwarded since then, such as those based on differential latencies or perceptual postdiction. In this chapter we aim to scrutinize the role of attention in either modulating or engendering the FLE.

23.1 Introduction

To deal with even simple challenges, such as grasping an object or avoiding a collision with either stationary or moving obstacles, everyday life demands from us the ability to localize a visual stimulus, within an acceptable degree of accuracy, in both space and time. Learning how to pin down the location of an object moving along its space–time trajectory in a given task depends on the one hand on the amount and quality of perceptual information provided by the sensory system, and on the other hand on the correctness of the action generated during that task. The behavioral outcome is continuously fed back to the nervous system, therefore constraining and refining, in an adaptive way, the representation of the world both in perception and in action.

However optimized our behavior turns out to be, the underlying perceptual edifice we assemble from the available sensory world is by no means unique. It is at most one efficient representation, among others, of the sensory world under circumstances that are changing. The multiplicity of perceptual solutions usually offered to a given sensory riddle lies at the core of the visual illusions, which we define as the mismatch between two or more percepts generated, under different conditions, by the same stimulus.

23.2 An illusion in space–time: the flash-lag effect

Until the nineteenth century, astronomers had to rely on the *eye and ear* method to gauge the space–time features of celestial events. For instance, they had to pinpoint the exact

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

moment at which a star crossed a marker in the telescope field by counting the ticks made by a clock while visually fixating the astronomical object. Recurring disagreement between the measurements made by different astronomers led, at that time, to the proposal that perceptual discrepancies might arise from the relative share of attention split between visual and auditory stimuli (Spence et al. 2001). In a more contemporary and prosaic scenario, assistant referees often make mistakes during soccer games when judging the relative location of an attacking player at the moment the ball touches or is passed by one of his or her teammates (Baldo, Ranvaud, & Morya, 2002).

The common root of the above examples is the need to perform a dual task: *Where* was the moving object (star or player) *when* the time marker was presented (a given tick or pass)? This question is intrinsically *spatial* and *temporal* in its very essence. In the midst of a variety of other visual illusions arising from the same question, the *flash-lag effect* (FLE) stands out as a still intriguing perceptual phenomenon, hotly debated over the last years. In the FLE, a stationary flash is perceived as spatially lagging a moving object, despite their being physically aligned to each other at that moment (Nijhawan 1994; Baldo & Klein 1995; Lappe & Krekelberg 1998; Purushothaman et al. 1998; Eagleman & Sejnowski 2000a; Nijhawan & Khurana 2000; Patel et al. 2000; Eagleman 2001; Krekelberg & Lappe 2001; Nijhawan 2002; Schlag & Schlag-Rey 2002; Whitney 2002; Ögmen et al. 2004; Baldo & Caticha 2005; Chappell et al. 2006; Nieman et al. 2006; Sarich et al. 2007; Linares & Lopez-Moliner 2007).

The FLE, in fact a centenarian illusion (Wundt 1874; Metzger 1931; Mackay 1958), was rediscovered by Romi Nijhawan (Nijhawan 1992, 1994). Nijhawan originally interpreted the flash-lag effect as resulting from a spatial extrapolation of the moving object ("motion extrapolation"), owing to the predictability of its trajectory. The perceptually extrapolated position of a moving object would thus compensate for the spatial error introduced by delays occurring throughout the visual system.

In 1995, we put forward the first challenge to Nijhawan's explanation by demonstrating that the magnitude of the FLE was dependent not only on the features of the moving object (such as its speed), as previously shown by Nijhawan, but on the visual eccentricity of the stationary flash as well (see Fig. 23.1(a)). This finding was at odds with the underpinnings of the motion extrapolation hypothesis. According to Nijhawan's proposal, the perceptual forward displacement of the moving object, in order to be adaptive, should depend on its kinematics features only, being independent of the psychophysical attributes of the flashing dot. However, the influence of the stationary object's physical features (such as its eccentricity and luminance) on the magnitude of the FLE has been consistently observed (Baldo & Klein 1995; Purushothaman et al. 1998; Patel et al. 2000; Baldo, Kihara, Namba et al. 2002).

We first interpreted the FLE "as resulting from a longer time delay involved in the visual processing of the flashing dots" (Baldo & Klein 1995). We proposed that some amount of time, dependent on eccentricity, is required to bring the flashing stimulus to a sufficiently high degree of perceptual processing for a snapshot of the moving stimulus to be taken. According to our original proposal, "such a time delay would be related to the abrupt

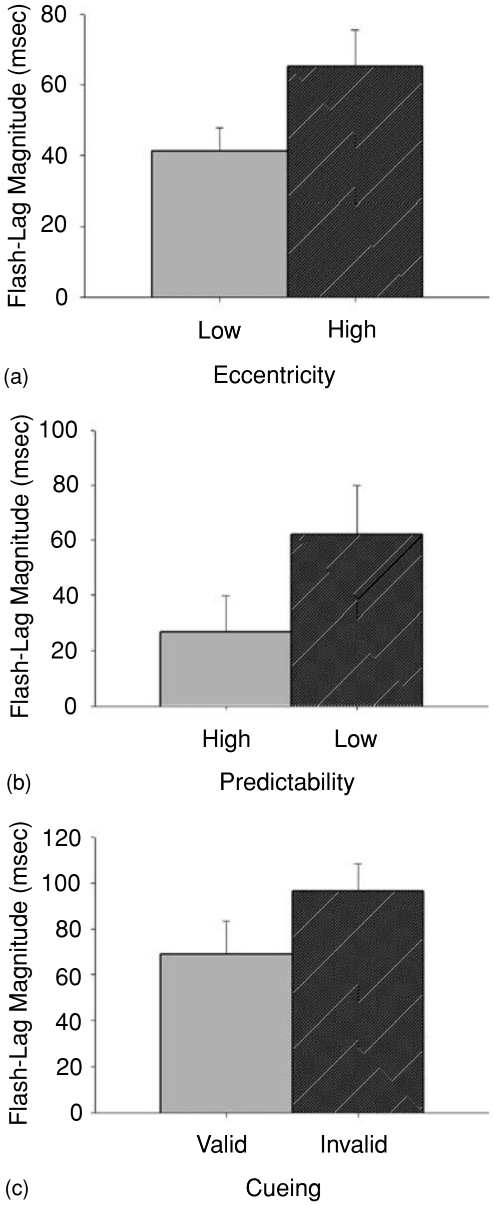


Fig. 23.1 Magnitude of the flash-lag effect obtained under different task conditions. (a) The visual eccentricity of the flashing stimulus has a significant effect on the magnitude of the perceptual phenomenon. “Low” and “High” correspond, respectively, to 1.7 and 3.9 deg of visual angle (Baldo, Kihara, Namba et al. 2002). (b) The magnitude of the flash-lag effect is also dependent on the spatial predictability of the flashing stimulus. “High” and “Low” stand for a predictability of 100% and 50%, respectively, regarding the presentation location of the flashing stimulus (Namba & Baldo 2004). (c) When attention is symbolically cued to the flashing stimulus’ potential location of appearance, the magnitude of the flash-lag effect varies according to the cueing validity. “Valid” indicates that the flashing stimulus indeed appears at the cued site; “Invalid” indicates otherwise (Namba & Baldo 2004).

onset of the flashing dots and might involve attentional mechanisms, either in capturing attention or in shifting the focus of attention from one place [the flashing dot] to another [the moving stimulus] across the visual field." Implicitly embedded in our account was the adoption of a spotlight metaphor as a framework to conceive the dynamics of visual attention (Tsal 1983).

Three years later, other independent groups reported on empirical findings that, yet again, could not be fitted into the motion extrapolation scheme. These authors showed, for instance, the robust dependence of the FLE on the trajectory of the moving stimulus (Lappe & Krekelberg 1998; Whitney & Murakami 1998) as well as on the luminance of both moving and stationary stimuli (Lappe & Krekelberg 1998; Purushothaman et al. 1998). From that time forth, the way was paved for a multitude of different approaches, views and accounts, within and around the flash-lag phenomenon (for reviews, see Nijhawan & Khurana 2000; Eagleman 2001; Krekelberg & Lappe 2001; Nijhawan 2002; Schlag & Schlag-Rey 2002; Whitney 2002).

23.3 The role of attention in the flash-lag effect

23.3.1 Does attention modulate the FLE?

Perception is strongly modulated by the organism's momentary goals and expectations. At least in part, this cognitive control over perception is due to attentional mechanisms. According to Palmer (1999), "[. . .] attention somehow plays a very important role in our conscious perception of visual events, by enabling non-conscious visual processing to reach consciousness and/or by creating durable representations in memory that can be used to report fleeting conscious perceptions that would otherwise disappear without a trace." As to the behavioral counterpart of visual attention, it manifests itself in speeding up reaction times and in increasing both sensitivity and efficiency of detection and identification of visual targets (Pashler 1998; Palmer 1999). These facilitatory processes may arise from either the goal-directed and voluntary (top-down) or stimulus-driven and automatic (bottom-up) deployment of attention.

Our first demonstration of a putative role of visual attention in modulating the flash-lag effect was based on a hazy empirical foundation (Baldo & Klein 1995). The main idea was centered on the time delay required to shift attention from the flashing stimulus' location to the moving stimulus' pathway, in order to make a spatial judgment of their relative locations. The influence of visual eccentricity on attentional shifts was, however, confounded with its even stronger influence on basic sensory issues, such as temporal latencies and spatial resolution. Later on, we were able to refine our analysis, providing more convincing support to the contribution of visual attention to the FLE (Fig. 23.1(b) and 23.1(c)). The magnitude of the FLE was shown to be modulated by the spatial predictability of the flashing stimulus (Baldo, Kihara, & Klein, 2000; Baldo, Kihara, Namba, et al. 2002; Baldo & Namba 2002) as well as symbolic cueing (Namba & Baldo 2004). Moreover, a modulatory role of either bottom-up or top-down deployment of attention was recently implemented in a simple feedforward neural network (Baldo & Caticha 2004, 2005; Cravo & Baldo 2008) that was able to not only replicate a long list of empirical findings but also encapsulate, in a unifying perspective, the main current accounts concerning the FLE.

Although challenged by conflicting results (Khurana & Nijhawan 1995; Khurana et al. 2000), the inherent involvement of attention in the FLE has been broadly supported by several other reports stemming from a diverse range of methodological and conceptual approaches (Eagleman & Sejnowski 2000b; Sheth et al. 2000; Rotman et al. 2002; Kanai et al. 2004; Vreven & Verghese 2005). More tellingly, recent findings by Chappell and colleagues have shown that the mere presence of a flash alters the perception of the position of a moving object by means of a stimulus-driven process they called *task-relevant attentional capture* (Chappell et al. 2006). Besides, by explicitly manipulating visual attention via a dual-task procedure, Sarich and colleagues were able to demonstrate the effect of dividing attentional resources on the flash-lag magnitude, thus strengthening the view according to which attentional processes should be included in a comprehensive model of the flash-lag illusion (Sarich et al. 2007).

The answer to this section's opening question, therefore, seems to be "yes, attention does have an effect on the FLE."

23.3.2 Does attention cause the FLE?

Considerable support from neuroanatomy and neurophysiology indicates that processing streams in the visual pathways are segregated so that perceptual dimensions such as location, motion, color, and object identity are processed in separate brain areas. Therefore, the attributes of an object, from local features to abstract properties, have to be bound together into a coherent representation by means of coordinated activity taking place across widespread neuronal populations. Also, because multiple objects fall within the usually large receptive field of a given neuron belonging to later stages of visual processing, stimuli have to compete to achieve perceptual representation. From empirical, computational, and conceptual grounds, a converging theoretical framework points to a fundamental role of attention in both processes: selection of competing stimuli and binding of multiple perceptual dimensions (Kastner & Ungerleider 2000; Engel et al. 2001; Corbetta & Shulman 2002; Lamme 2003; Dehaene et al. 2006; Roelfsema 2006; Serences & Yantis 2006).

Everyone who was ever engaged in judging, during a psychophysical procedure, the relative locations of moving and flashing stimuli knows how effortful the process of focusing attention on the perceptual task is. In fact, a whole set of cognitive operations is in progress during even such a simple perceptual task, where top-down factors include goal definition, action planning, working memory, and selective attention (Kastner & Ungerleider 2000; Engel et al. 2001; Corbetta & Shulman 2002).

In Schlag's and Schlag-Rey's own words, "once the signal of an event is received by the brain, it can be used for sampling, probing or starting a process" (Schlag & Schlag-Rey 2002). An abrupt-onset event (flash) would thus start the process of the perceptual comparison between the moving and flashing stimuli's relative locations. This process should require an attention shift from the flashing stimulus to the moving stimulus and would, therefore, take some amount of time. Although quite similar to our original account regarding the involvement of attention in the FLE, the above statement differs in an important way. The shift from the stationary stimulus to the moving one does not have to occur within physical space, but it might be an object-based shift instead.

Growing evidence, from both psychophysical and neurophysiological approaches, has shown that selective attention operates not only on a space-based representation but undergoes object-based reallocations in which the boundaries of extended objects also determine what is attended and how attention is deployed (Corbetta & Shulman 2002; Yantis & Serences 2003; Khayat et al. 2006). Therefore, attending to a spatial location is just one way the brain can select relevant sensory information. Attention can also be directed to different features of an object (such as its color and shape) or can even be shifted between different objects. A frontoparietal cortical network known to be recruited for spatial attention also seems to be involved in other kinds of visual selection (Corbetta & Shulman 2002; Lamme 2003).

In view of the fact that the FLE endures even under circumstances in which there is no sizeable spatial separation between the flashed and the moving stimuli (Khurana & Nijhawan 1995), a spatial shift of visual attention needn't play a relevant part in generating this perceptual illusion. However, we can conceive that, starting with the detection of the flash, a shift (or spread) of attention from the stationary *object* to the moving *object* has to occur to bind them together into a unitary package, allowing the completion of the perceptual judgment. This object-based attentional shift would most likely require time to be carried out, regardless of any spatial separation between both visual stimuli. The time spent shifting attention between different objects (or, alternatively, in spreading over different parts of a single object) would lead to a percept whose components consist of a moving stimulus further ahead along its motion pathway. Of course, postulating an attentional shift from the stationary to the moving object assumes that attention should be initially grabbed by the abrupt-onset stationary stimulus. In fact, the relevance of this stimulus-driven attentional capture was demonstrated in the context of the FLE (Chappell et al. 2006). Yet we can also consider a nondirectional attentional spread over the object (in feature space) that starts *when* the abrupt-onset stimulus is presented but not necessarily from *where* it is presented. In this sense, "attention" would be a time-consuming process required to bind together the parts of the object into a unified and coherent whole (where the time elapsed would, again, lead to the flash-lag effect). The idea of a time-consuming process appears also under different names in different frameworks, for instance the "postdiction" (Eagleman & Sejnowski 2000a) and "sampling" (Brenner & Smeets 2000) hypotheses, which nonetheless may share important underlying aspects.

Temporal delays associated with the integration of a spatially extended object have been linked by other scholars as well to a time-consuming shift or spread of visual attention, concerning both visual perception in general (Roelfsema et al. 2000; Houtkamp et al. 2003; Enns & Oriet 2004) and the FLE in particular (Kanai et al. 2004; Chappell et al. 2006; Sarich et al. 2007). Also, the central idea presented here is akin to the "object updating" scheme offered by Enns and Oriet (2004) in order to account for the perceptual asynchrony illusion, PAI (Moutoussis & Zeki 1997). In the PAI, observers are asked to report one attribute (e.g., color: red vs. blue) that is cyclically alternating with another, also alternating, different attribute (e.g., orientation: vertical vs. horizontal). Enns and Oriet (2004) interpret the observed perceptual asynchrony as the consequence of shifting attention from a "defining" attribute (say, color) to a "report" attribute (say, orientation). Enns and Oriet (2004), to the extent of our knowledge, were the first authors to point out a functional connection between the PAI and the FLE, widening the focus of empirical and conceptual approaches. The temporal continuity of an object as such seems also to be important to preserve the flash-lag illusion. Moore and Enns (2004) have shown that disrupting the continuity of the moving object by a sudden change in color or size diminishes or even abolishes the FLE. This phenomenon, although still compatible with the present object-based attentional account, calls for a more elaborate model of visual integration, because attention is now being grabbed by another transient stimulus, which leads to a newly formed object that must also compete for binding.

A neat consequence of the present account relies on its generality. Because the main idea pivots on an attentional shift (or attentional binding) in *feature space* rather than *physical space*, the engendering mechanism would hold as long as a perceptual comparison is to be made between a reference, "stationary," stimulus and a "changing" stimulus, where "stationary" and "changing" refer to variables in any sensory feature space. A sole, object-based attentional shift or spread mechanism could, therefore, account for the appearance of the FLE in tasks, wherein the perceptually assessed variables go beyond spatial localization and comprise features such as color, luminance, and even stimulus complexity (Sheth et al. 2000). Still more general is the fact that the present attentional account conveys the potential emergence of the FLE in a sensory modality other than the visual system, or even across different modalities (Alais & Burr 2003). Under these broad perceptual scenarios, the recipe we offer is simple: given that a reference, abrupt-onset, stimulus is perceived, attention is recruited to bind it to a changing stimulus, whose current status must be recorded; because this attentional binding takes some amount of time, the changing variable will be pinned down at an advanced location within its feature space.

In an attempt to sum up, one possible answer to this section's question is: "Yes, attentional mechanisms might well be responsible for causing the FLE."

23.4 Epilogue

Over the years, several reports have resorted to attentional mechanisms to explain a class of perceptual phenomena closely related to FLE. The relationship between the FLE and those phenomena relies on the very nature of the perceptual judgment: a visual localization in space and time. Among these phenomena, one of the oldest is the Fröhlich effect, in which a moving object coming out from an occluding edge is seen to first appear at some distance from the edge (Fröhlich 1923; Müsseler & Aschersleben 1998; Kirschfeld & Kammer 1999; Kerzel & Müsseler 2002; Kerzel & Gegenfurtner 2004). Attention-shifting explanations have been offered as the basis of the Fröhlich effect (Müsseler & Neumann 1992; Müsseler & Aschersleben 1998; Kirschfeld & Kammer 1999; Kerzel & Müsseler 2002; Kerzel & Gegenfurtner 2004) as well as of several other visual illusions, such as the representational momentum (Hayes & Freyd 2002; Kerzel 2003; Hubbard 2005), the line motion illusion (Hikosaka et al. 1993; Kirschfeld & Kammer 2000), the onset repulsion effect (Kerzel & Gegenfurtner 2004), the tandem effect (Müsseler & Neumann 1992), and the flash-shift effect (Shim & Cavanagh 2003).

Yet, we have to be careful not to convert attentional explanations into a theoretical panacea for the empirical queries we struggle to solve. Given the misty state of affairs in which conceptual and experimental models of attention unfold, we have to be fully aware of the dangers of straightforwardly invoking a cure-all wonder solution ready to crack a stiff cluster of problems. Otherwise, we can quickly come up with an explanation that either lacks a theoretical stem deeply rooted in psychophysical and neurophysiological foundations or lacks empirical content, being shielded against experimental refutation.

As a final question: Is the FLE a sensory phenomenon modulated by attentional mechanisms or an attentional phenomenon modulated by sensory mechanisms? Indisputably, not only the FLE but also the building up of any percept depends on the stimulus parameters. Therefore, the time-consuming process involved in attentionally binding multiple objects (or a single extended object) must be added up to the temporal delays brought about by the basic neural mechanisms operating along the entire sensory pathways. The FLE, as a spatiotemporal illusion, might encompass spatial and temporal mechanisms ranging from basic and local neuronal interactions (Berry et al. 1999; Erlhagen 2003; Cantor & Schor 2004; Baldo & Caticha 2005; Cravo & Baldo 2008) to cognitive integration over distributed cortical networks (such as discussed in the present chapter). For that reason, we should not be surprised by observing such an enormous susceptibility in the perceptual outcome involving the FLE. On the sensory side, a flash-lag may even be turned into a flash-lead effect under the appropriate luminance settings (Purushothaman et al. 1998; Patel et al. 2000); on the cognitive side, the magnitude of FLE may be reduced by disrupting the temporal continuity of the moving object (Moore & Enns 2004) or inflated by withdrawing attentional resources from the target (Sarich et al. 2007).

In conclusion, we believe that the flash-lag phenomenon arises from a spatiotemporal integration that involves a time-consuming cognitive process, most likely an attentional mechanism required to select and bind perceptual features distributed over widespread neural networks. Naturally, psychophysical parameters also shape the dynamics of the spatial and temporal processing, being an integral part of the overall phenomenon. Therefore, questions such as those in the previous paragraph seem to be misleading. Both sensory and attentional mechanisms may be considered either causal or modulatory factors depending, for instance, on the specific experimental strategy devised to approach the flash-lag phenomenon. In view of that, a fruitful investigative program should rely not only on an unremitting search for relevant experimental findings but also on a deeper scrutiny of the words and concepts we choose to explain these findings.

Acknowledgments

We thank Peter Claessens, Milene Regatão, Hamilton Haddad, Chris Cantor, Hulusi Kafaligonul and an anonymous reviewer for helpful comments and suggestions. This research has been supported in part by FAPESP and CNPq.

References

Alais, D., & Burr, D. (2003). The flash-lag effect occurs in audition and cross-modally. *Curr Biol* **13**: 59–63.

Baldo, M. V. C., & Caticha, N. (2004). The flash-lag and Fröhlich effects caught by the net: computational modeling of visual illusions. In A. M. Oliveira, M. Teixeira, G. F. Borges, & M. J. Ferro (eds.), *Annual Meeting of the International Society for Psychophysics* (Coimbra, Portugal) 222–227.

Baldo, M. V. C., & Caticha, N. (2005). Computational neurobiology of the flash-lag effect. *Vision Res* **45**: 2620–2630.

Baldo, M. V. C., Kihara, A. H., & Klein, S. A. (2000). Lagging behind because of sensory and attentional delays. [Abstract] *Invest Ophthalmol Vis Sci* **41**: S420.

Baldo, M. V. C., Kihara, A. H., Namba, J., & Klein, S. A. (2002). Evidence for an attentional component of perceptual misalignment between moving and flashing stimuli. *Perception* **31**: 17–30.

Baldo, M. V. C., & Klein, S. A. (1995). Extrapolation or attention shift? *Nature* **378**: 565–566.

Baldo, M. V. C., & Namba, J. (2002). The attentional modulation of the flash-lag effect. *Braz J Med Biol Res* **35**: 969–972.

Baldo, M. V. C., Ranvaud, R. D., & Morya, E. (2002). Flag errors in soccer games: the flash-lag effect brought to real life. *Perception* **31**: 1205–1210.

Berry, M. J., Brivanlou, I. H., Jordan, T. A., & Meister, M. (1999). Anticipation of moving stimuli by the retina. *Nature* **398**: 334–338.

Brenner, E., & Smeets, J. B. J. (2000). Computational neurobiology of the flash-lag effect. *Vision Res* **40**: 1645–1648.

Cantor, C. R. L., & Schor, C. M. (2004). Does the temporal impulse response cause the flash-lag effect? [Abstract] *J Vis* **4**(8): 72, 72a, http://journalofvision.org/4/8/72/, doi: 10.1167/4.8.72.

Chappell, M., Hine, T. J., Acworth, C., & Hardwick, D. R. (2006). Attention "capture" by the flash-lag flash. *Vision Res* **46**: 3205–3213.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci* **3**: 201–215

Cravo, A. M., & Baldo, M. V. C. (2008). A psychophysical and computational analysis of the spatio-temporal mechanisms underlying the flash-lag effect. *Perception* **37**: 1850–1866.

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends Cogn Sci* **10**: 204–211.

Eagleman, D. M. (2001). Visual illusions and neurobiology. *Nat Rev Neurosci* **2**: 920–926.

Eagleman, D. M., & Sejnowski, T. J. (2000a). Motion integration and postdiction in visual awareness. *Science* **287**: 2036–2038.

Eagleman, D. M., & Sejnowski, T. J. (2000b). Reply to Krekelberg et al. *Science* **289** (5482): 1107a.

Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top-down processing. *Nat Rev Neurosci* **2**: 704–716.

Enns, J. T., & Oriet, C. (2004). Perceptual asynchrony: modularity of consciousness or object updating? [Abstract] *J Vis* **4**(8): **27**, 27a, http://journalofvision.org/4/8/27/, doi: 10.1167/4.8.27.

Erlhagen, W. (2003). Internal models for visual perception. *Biol Cybern* **88**: 409–417.

Fröhlich, F. W. (1923). Über die Messung der Empfindungszeit. *Zeitschrift für Sinnesphysiologie* **54**: 58–78.

Hayes, A. E., & Freyd, J. J. (2002). Representational momentum when attention is divided. *Vis Cogn* **9**: 8–27.

Hikosaka, O., Miyauchi, S., & Shimojo, S. (1993). Focal visual attention produces illusory temporal order and motion sensation. *Vision Research* **33**: 1219–1240.

Houtkamp, R., Spekreijse, H., & Roelfsema, P. R. (2003). A gradual spread of attention during mental curve tracing. *Perception & Psychophysics* **65**: 1136–1144.

Hubbard, T. L. (2005). Representational momentum and related displacements in spatial memory: A review of the findings. *Psychonom Bull Rev* **12**: 822–851.

Kanai, R., Sheth, B. R., & Shimojo, S. (2004). Stopping the motion and sleuthing the flash-lag effect: spatial uncertainty is the key to perceptual mislocalization. *Vision Res* **44**: 2605–2619.

Kastner, S., & Ungerleider, L. G. (2000). Mechanisms of visual attention in the human cortex. *Ann Rev Neurosci* **23**: 315–341.

Kerzel, D. (2003). Attention maintains mental extrapolation of target position: irrelevant distractors eliminate forward displacement after implied motion. *Cognition* **88**: 109–131.

Kerzel, D., & Gegenfurtner, K. R. (2004). Spatial distortions and processing latencies in the onset repulsion and Fröhlich effects. *Vision Res* **44**: 577–590.

Kerzel, D., & Müsseler, J. (2002). Effects of stimulus material on the Fröhlich illusion. *Vision Res* **42**: 181–189.

Khayat, P. S., Spekreijse, H., & Roelfsema, P. R. (2006). Attention lights up new object representations before the old ones fade away. *J Neurosci* **26**: 138–142.

Khurana, B., & Nijhawan, R. (1995). Reply to Baldo and Klein. *Nature* **378**: 566.

Khurana, B., Watanabe, K., & Nijhawan, R. (2000). The role of attention in motion extrapolation: Are moving objects "corrected" or flashed objects attentionally delayed? *Perception* **29**: 675–692.

Kirschfeld, K., & Kammer, T. (1999). The Fröhlich effect: a consequence of the interaction of visual focal attention and metacontrast. *Vision Res* **39**: 3702–3709.

Kirschfeld, K., & Kammer, T. (2000). Visual attention and metacontrast modify latency to perception in opposite directions. *Vision Res* **40**: 1027–1033.

Krekelberg, B., & Lappe, M. (2001). Neuronal latencies and the position of moving objects. *Trends Neurosci* **24**: 335–339.

Lamme, V. A. F. (2003). Why visual attention and awareness are different. *Trends Cogn Sci* **7**: 12–18.

Lappe, M., & Krekelberg, B. (1998). The position of moving objects. *Perception* **27**: 1437–1449.

Linares, D., & Lopez-Moliner, J. (2007). Absence of flash-lag when judging global shape from local positions. *Vision Res* **47**: 357–362.

Mackay, D. M. (1958). Perceptual stability of a stroboscopically lit visual field containing selfluminous objects. *Nature* **181**: 507–508.

Metzger, W. (1931). Versuch einer gemeinsamen Theorie der Phänomene Fröhlichs und Hazelhoffs und Kritik ihrer Verfahren zur Messung der Empfindungszeit. *Psychologische Forschung* **16**: 176–200.

Moore, C. M., & Enns, J. T. (2004). Object updating and the flash-lag effect. *Psychol Sci* **15**: 866–871.

Müsseler, J., & Aschersleben, G. (1998). Localizing the first position of a moving stimulus: the Fröhlich effect and attention shifting explanation. *Perception and Psychophysics* **60**: 683–695.

Müsseler, J., & Neumann, O. (1992). Apparent distance reduction with moving stimuli (tandem effect) – Evidence for an attention-shifting model. *Psychol Res-Psychologische Forschung* **54**: 246–266.

Namba, J., & Baldo, M. V. C. (2004). The modulation of the flash-lag effect by voluntary attention. *Perception* **34**: 621–631.

Nieman, D., Nijhawan, R., Khurana, B., & Shimojo, S. (2006). Cyclopean flash-lag illusion. *Vision Res* **46**: 3909–3914.

Nijhawan, R. (1992). Misalignment of contours through the interaction of apparent and real motion systems. [Abstract] *Invest Ophthalmol Vis Sci* **33**: 1415.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature* **370**: 256–257.

Nijhawan, R. (2002). Neural delays, visual motion and the flash-lag effect. *Trends Cogn Sci* **6**: 387–393.

Nijhawan, R., & Khurana, B. (2000). Conscious registration of continuous and discrete visual events. In T. Metzinger (ed.), *Neural Correlates of Consciousness: Empirical and Conceptual Questions* (203–219). Cambridge, MA: MIT Press.

Öğmen, H., Patel, S. S., Bedell, H. E., & Camuz, K. (2004). Differential latencies and the dynamics of the position computation process for moving targets, assessed with the flash-lag effect. *Vision Res* **44**: 2109–2128.

Palmer, S. E. (1999). *Vision Science: Photons to Phenomenology*. Cambridge, MA: MIT Press.

Pashler, H. E. (1998). *The Psychology of Attention*. Cambridge, MA: MIT Press.

Patel, S. S., Öğmen, H., Bedell, H. E., & Sampath, V. (2000). Flash-lag effect: differential latency, not postdiction. *Science* **290**: 1051a.

Purushothaman, G., Patel, S. S., Bedell, H. E., & Öğmen, H. (1998). Moving ahead through differential visual latency. *Nature* **396**: 424.

Roelfsema, P. R. (2006). Cortical algorithms for perceptual grouping. *Ann Rev Neurosci* **29**: 203–227.

Roelfsema, P. R., Lamme, V. A. F., & Spekreijse, H. (2000). The implementation of visual routines. *Vision Res* **40**: 1385–1411.

Rotman, G., Brenner, E., & Smeets, J. B. J. (2002). Spatial but not temporal cueing influences the mislocalisation of a target flashed during smooth pursuit. *Perception* **31**: 1195–1203.

Sarich, D., Chappell, M., & Burgess, C. (2007). Dividing attention in the flash-lag illusion. *Vision Res* **47**: 544–547.

Schlag, J., & Schlag-Rey, M. (2002). Through the eye, slowly: delays and localization errors in the visual system. *Nat Rev Neurosci* **3**: 191–200.

Serences, J. T., & Yantis, S. (2006). Selective visual attention and perceptual coherence. *Trends Cogn Sci* **10**: 38–45.

Sheth, B., Nijhawan, R., & Shimojo, S. (2000). Changing objects lead briefly flashed ones. *Nat Neurosci* **3**: 489–495.

Shim, W. M., & Cavanagh, P. (2003). Attentive tracking can modulate the illusory misalignment of a flash. [Abstract] *J Vis* **3**(9): 188a, http://journalofvision.org/3/9/188/, doi: 10.1167/3.9.188

Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *J Exp Psychol* **4**: 799–832.

Tsal, Y. (1983). Movements of attention across the visual field. *J Exp Psychol Hum Percept Perform* **9**: 523–530.

Vreven, D., & Verghese, P. (2005). Predictability and the dynamics of position processing in the flash-lag effect. *Perception* **34**: 31–44.

Whitney, D. (2002). The influence of visual motion on perceived position. *Trends Cogn Sci* **6**: 211–216.

Whitney, D., & Murakami, I. (1998). Latency difference not spatial extrapolation. *Nat Neurosci* **1**: 656–657.

Wundt, W. (1874). *Grundzüge der physiologischen psychologies*. Leipzig, Germany: Wilhelm Engelmann.

Yantis, S., & Serences, J. T. (2003). Cortical mechanisms of space-based and object-based attentional control. *Curr Opin Neurobiol* **13**: 187–193.

24

Illusions of time, space, and motion: flash-lag meets chopsticks and reversed phi

STUART ANSTIS

Summary

In the "chopstick illusion" (Anstis 1990, 2003) a vertical and horizontal line overlapped to form a cross and followed clockwise circular orbits in counterphase, with one line being at 6 o'clock when the other was at 12 o'clock. The intersection of the lines moved counterclockwise, but it was wrongly perceived as rotating clockwise. This chopstick illusion reveals how moving objects are parsed, based upon the intrinsic and extrinsic terminators of lines viewed through apertures. We conclude that intersections were not parsed as objects, but instead the motion of the terminators (tips) propagated along the lines and was blindly assigned to the intersection. In the similar "sliding rings illusion," we found that observers could use their eyes to track intersections only when these appeared rigid and not when they appeared to slide. Conclusion – smooth pursuit eye movements are under top-down control and are compelled to rely upon perceptual interpretation of objects.

In the "flash-lag" effect, a static object that is briefly flashed up next to a moving object appears to lag behind the moving object (Nijhawan 2002). We superimposed a flashed spot on a chopsticks intersection that appeared to be moving clockwise along a circular path but was actually moving counterclockwise. We found that the flash appeared displaced clockwise. This was appropriate to the physical, not the subjective direction of rotation, indicating that the flash-lag and the chopstick illusions coexist without interacting. Similarly, the flash-lag effect was unaffected by reversed phi. Probably the flash-lag occurs early in the visual system, before motion signals are parsed into moving objects.

24.1 Introduction

In 1990, I published a "chopstick illusion" (Anstis 1990), as shown in Fig. 24.1(a). A vertical line and a superimposed horizontal line move in counterphase along clockwise circular paths, without rotating. The chopstick illusion arises in the central intersection, where the two lines cross. This actually moves counterclockwise around a circle, but it is incorrectly perceived as moving clockwise. In our view, the local motion signals from the intersection are ignored because the sliding intersection is not parsed as an object. Instead, clockwise motion signals from the terminators propagate along the two lines and are blindly assigned to the central intersection – even though there is no single spatial phase that can be assigned to this intersection.

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

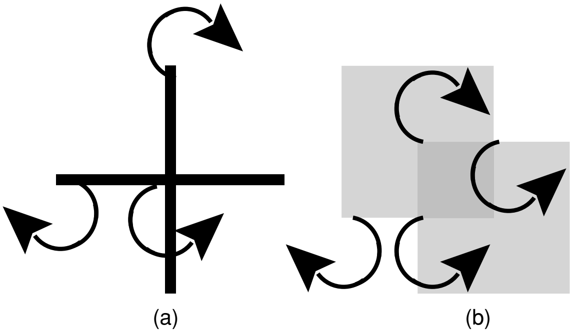


Fig. 24.1 The chopstick illusion.

24.2 Observations from naive students

A class of 208 undergraduate students in a large lecture hall viewed versions of the chopstick illusion from a great variety of viewing distances and angles. Results were consistent across observers. First of all, 99% of them correctly identified the path of a control stimulus – a small plus sign that circled clockwise. The plus sign was then replaced with a large cross whose horizontal and vertical arms ran off the side and top and bottom edges of the screen. Surprisingly, only 54% saw it correctly as clockwise, 46% incorrectly as counterclockwise. These close-to-chance results suggest that line terminators contribute much more than the central intersection toward judgments of trajectories. When shown a version of Fig. 24.1(a), in which the lines moved clockwise and the central intersection moved counterclockwise, 97% of the observers incorrectly reported the central intersection as moving clockwise. Thus almost everybody experienced the chopstick illusion.

24.2.1 Transparent edges

In Fig. 24.1, edges work just as well as lines. In Fig. 24.1(b), two overlapping transparent squares both move clockwise. Their two intersection points actually move counterclockwise but appear to move clockwise. 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).

24.2.2 Aperture effects

If Fig. 24.1 is viewed through a stationary square aperture (Fig. 24.2), the illusion disappears and the central intersection is perceived as a single rigid cross circling counterclockwise instead of as two sliding lines circling clockwise. The square aperture hides the clockwise signals that usually come from the ends of the lines, but in Fig. 24.2(b) the floating outline just touches all four tips of the two rotating lines and the lines can still be seen in their entirety, so all the terminator information is still available. Yet now, the illusion is gone and

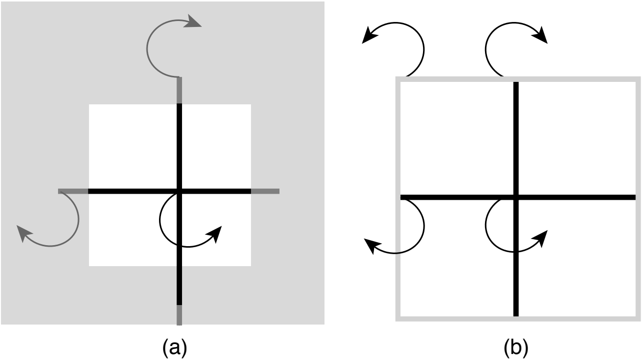


Fig. 24.2 The chopstick stimulus seen through a stationary aperture.

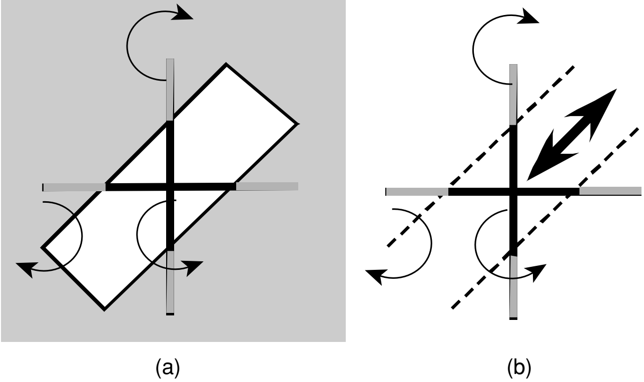


Fig. 24.3 (a) The chopstick stimulus seen through an oblique slot in a black mask. (b) Same as (a), except the mask is white and invisible.

the contrast intersection is correctly seen as a rigid cross moving counterclockwise. Why should the illusion vanish when the entire rods are still visible? Probably, the bare line ends in Fig. 24.1 are taken as "intrinsic terminators" (Shimojo et al. 1989) – the actual ends of the lines – whereas when the very same ends touch the floating frame in Fig. 24.2(b), they are interpreted as "extrinsic terminators," hidden behind the square frame and giving no clues as to the real motion of the lines. So the visual system perceives a rigid, not a sliding cross, whose motion is correctly determined locally as being counterclockwise.

Figure 24.3(a) shows the chopstick stimulus of Fig. 24.1 viewed through an oblique slot in a *visible* black mask. Result: 87% of observers perceived this as a rigid cross circling coherently counterclockwise. Here and in Fig. 24.2(a), the ends of the lines were perceived as "extrinsic" (Shimojo et al. 1989), that is, as occluded by the aperture and extending behind it, and did not influence the perceived motion of the central intersection.

Figure 24.3(b) is *identical* to Fig. 24.3(a), except that now the mask is white and *invisible*. Result: 94% of observers perceived the intersection as sliding along an oblique axis parallel

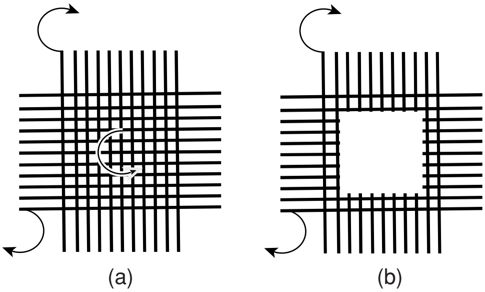

(a) (b)

Fig. 24.4 (a) The chopstick stimulus made of twenty-two horizontal and twenty-two vertical lines. Reduced number of lines is shown for clarity. (b) The same stimulus as (a), except with a hole in the center.

to the slot (thick oblique arrow). Here the line tips were perceived as real or "intrinsic" and their oblique motion determined the perceived motion of the center. The individual lines were perceived as moving obliquely and the central intersection likewise. Thus parsing of the terminators radically altered both perceived rigidity and perceived direction of motion.

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."

24.2.3 Multiple lines

In Fig. 24.1, 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 Fig. 24.4, there are 22 vertical and 22 horizontal lines, which gives 88 terminators and 484 (22^2) intersections. Now the intersections outnumber the terminators by 5.5:1, yet a chopstick illusion is still seen. We conclude that the chopstick illusion is not caused by mere force of numbers. (For clarity the number of lines is reduced in Fig. 24.4.)

We wondered whether the terminators won out over the intersections merely because the terminators were on the outside of the figure. We can take as an analogy Sinha and Poggio's (1996) well-known two faces of Clinton side by side. One had Clinton's hair and the other had Gore's hair (superimposed by means of Photoshop). Naive observers identified the two men as Clinton and Gore, showing that they used the outer hairline rather than internal features (eyes, nose, mouth) to identify the faces. Perhaps the movement of a display tends to be identified by the movement of its outermost parts? We falsified this by punching a hole in the middle of Fig. 24.4(a). Now, as before, the terminators moved clockwise and the intersections still moved counterclockwise but appeared to move clockwise in the

Fig. 24.5 The chopstick stimulus with *stationary* texture (printed text).

usual chopstick illusion. But now, a central white square was visible, defined by the inner terminators around the edges of the "hole" (Fig. 24.4(b)). This hole was correctly seen as moving clockwise. We conclude that terminators drove the perception of motion in this display, no matter whether their position was internal or external and no matter whether they pointed inward or outward.

24.2.4 Texture

Information that a line moves along its own length can come from the line terminators or equally well from texture painted on the line. But in Fig. 24.5 we gated *stationary* texture (printed text) with moving lines. Result: The stimulus underwent scission or splitting into different perceptual depth planes, with the lines looking like transparent slits or elongated portholes through which slices of a stationary background texture could be seen. The clockwise chopstick illusion was still present, so the clockwise terminators won out over the stationary dots and over the counterclockwise intersection.

24.2.5 Tip paths matter, intersection's path does not

The circular path of the central intersection is geometrically a Lissajous figure (Jean Antoine Lissajous 1822–1880). The horizontal position x of the intersection is determined by the vertical rod, and the vertical position of the intersection is determined by the horizontal rod: $x = r * \sin(a)$, $y = r * \cos(a)$. Although the intersection rotates clockwise, it appears to share the counterclockwise motion of the tips. Suppose we double the speed of the horizontal rod but do not change the vertical rod, so that $x = r * \sin(a)$, $y = r * \cos(2 * a)$. Now the intersection's path is a vertical figure 8 – also a Lissajou figure. However, observers still perceive the path of the intersection to be a clockwise circle – they cannot see the figure 8 at all. Finally, let us make the vertical rod trace out a vertical figure 8 while the horizontal rod traces out a horizontal figure 8 (Fig. 24.6). Now the intersection appears to trace out a "crazy" path, like some combination of figure 8s that the observers are unable

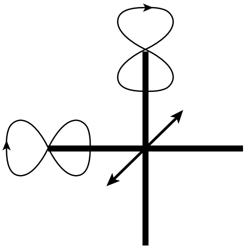


Fig. 24.6 Each line moves along a figure 8 path. The intersection actually moves obliquely (central arrow), but observers perceive it as following a complicated crazy path.

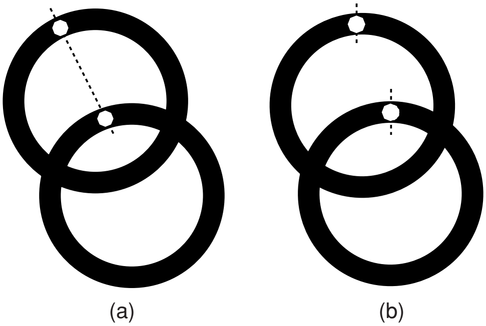


Fig. 24.7 Ambiguously moving rings. (a) Rings with painted-on dots appear to rotate as a single rigid figure 8. (b) When the dots float on the rings, always remaining at 12 o'clock, the two rings appear to rotate independently and slide over each other.

to resolve. If asked to sketch the intersection's path on a piece of paper, they usually declare that they cannot. Yet if an opaque screen with a circular hole is dropped over the display, they see at once that the intersection is merely moving back and forth along an oblique line! This simple path is completely unavailable to conscious scrutiny. Thus covering up the line tips allows observers to see the true motion of the intersection, and leaving them visible masks this motion. It is clear that the perceived path of the intersection is strongly driven by the line terminators, and only very weakly, if at all, by the actual path of the intersection.

24.2.6 Sliding rings

In Fig. 24.7 the terminators were removed by bending the lines around into rings. One dot painted on each ring radically altered the perceived motion. When the dots rotated in step with the rings (Fig. 24.7(a)) observers reported a rigid welded figure 8, rotating coherently. This satisfies the rigidity constraint (Ullman 1979). When the dots floated at 12 o'clock on each ring (Fig. 24.7(b)) observers reported two separate rings sliding over each other. This minimizes the motion within each ring by sacrificing rigidity (Ullman 1979; Shiffrar & Pavel 1991). Thus, each ring coheres with its dot rather than with its intersection. These results cannot be predicted from the vague idea that the visual system prefers "simplicity" or a "good Gestalt."

A naive observer tracked the intersection of the two rings in Fig. 24.7(a) or (b) while his eye movements were recorded. The mean tracking errors for the intersections of the rigid figure 8 (a) and of the two sliding rings (b) were, respectively, 1.04 deg and 9.93 deg of visual angle (Shimozaki & Ballard, unpublished data). This almost tenfold ratio in tracking errors suggests that smooth pursuit movements are not merely a bottom-up retinal feedback system (Lisberger et al. 1987; Krauzlis 1994) but may be strongly influenced by top-down cognitive processes such as object interpretation (Kowler 1990).

24.3 Chopsticks and the flash-lag effect

In the well-known "flash-lag" illusion (Mackay 1961), a flashed stimulus is presented physically aligned with a continuously moving object, and the flash is visible in a lagging position relative to the moving object (reviewed by Nijhawan 2002). This illusion has been variously attributed to motion extrapolation (Khurana et al. 2000), in which the observer predicts where a moving object will be soon, rather like a hunter who shoots ahead of a rabbit instead of directly at it. It has been attributed to differential latency (Whitney, Murakami, et al. 2000; Whitney, Cavanagh, et al. 2000), in which the neural signal from a flashed object is thought to be delayed relative to the signal from a moving object. Eagleman and Sejnowski (2000) attributed it not to prediction but to postdiction. Krekelberg and Lappe (2000) suggest that the perceived distance between a moving and a flashed dot is not determined from a "snapshot" of the image but by temporal averaging. The distance is averaged over a time period of about 600 msec, which would allow the visual system to improve its accuracy but would also lead to a misperception of the distance between the objects. Finally, Baldo et al. (2002) attribute the effect to attentional allocation.

We combined the flash-lag effect with the chopstick illusion. The brief flash was superimposed on the chopstick intersection in Fig. 24.1(a) – an intersection that was really moving to the left but apparently moving to the right. We tested whether any flash-lag effect would be appropriate to the physical or perceived direction of rotation. A flash that lagged behind the true physical rotation would suggest that the flash-lag effect occurred early in the visual system, before local motion signals were parsed into moving objects. On the other hand, a flash that lagged behind the illusory direction of rotation would suggest that the visual system parsed the motion first and saw the flash-lag illusion second.

Method. Three observers viewed a version of Fig. 24.1(a), in which the intersection followed a circular path of diameter 6 deg at a rotation rate of 0.9 rev/sec. When each line moved clockwise, their central intersection followed a path that also appeared to be clockwise but was actually counterclockwise. Once on every rotation a small white disk was flashed up exactly centered in the intersection. Observers gazed at a central fixation point and were asked to report upon the perceived position of the flashed disk relative to the intersection. The flash could occur at any of eight positions around the trajectory. The flash was centered on the moving intersection, but the flash-lag illusion shifted its

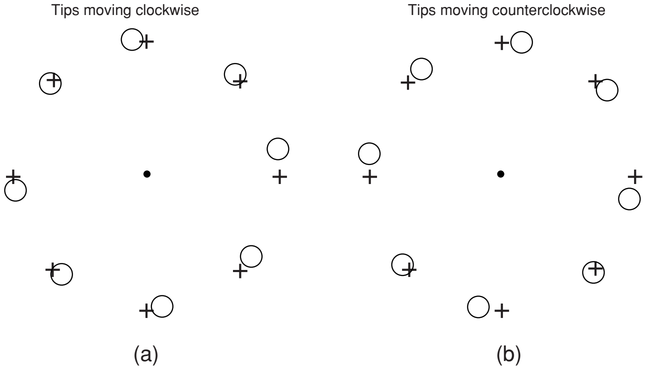


Fig. 24.8 Results of a flash-lag experiment in which a flash was superimposed at eight different positions of a rotating intersection in the chopstick illusion for (a) intersections moving counterclockwise, (b) intersections moving clockwise. Crosses show positions of intersections and disks show the objective nulling position of the flash that appeared subjectively superimposed on the intersections. Results are appropriate to physical, not illusory, direction of movement of the intersection of the lines.

apparent position, either clockwise or counterclockwise. The observer nulled out this flash-lag illusion by striking one of two keys that physically offset the flash either clockwise or counterclockwise. This titrated the illusion against a compensatory spatial offset, which was recorded for later analysis (Anstis 2007).

Results. These are shown in Fig. 24.8, with results for clockwise and counterclockwise rotations being plotted separately (mean of three observers). Figure 24.8(a) shows that when the intersections of the rods moved counterclockwise, the null position of the flash, at which it appeared superimposed on the intersection, was shifted counterclockwise through a mean angular rotation of 6.7 deg. So without this nulling offset, each flash would have appeared in a clockwise – shifted position, where the moving intersection had just been some 20.8 msec ago. This clockwise subjective lag, caused by the flash-lag effect, is appropriate to the *physical* counterclockwise motion of the intersection, *not* to its subjective clockwise motion. Correspondingly, when the intersections of the rods moved clockwise in Fig. 24.8(b), the nulling offset of the flash was shifted clockwise by 19 msec.

We conclude that the direction of the flash-lag effect was consistent with the objective, not the subjective motion of the intersection. Therefore the flash-lag effect was not influenced by the chopstick illusion, which implies that it occurs early in the visual system, *before* local motion signals are processed to give parsed moving objects.

Were these results obvious? After I had collected the data, I informally approached two leading authorities, named A and B, who have both published on the flash-lag effect. I asked them to guess whether I had found that the flash-lag effect was (1) absent, (2) appropriate to the true, physical direction of rotation, or (3) appropriate to the illusory direction of rotation. Chance performance would be 1.5 guesses, on average. A took two guesses and B took three. Names are withheld by request of A and B.

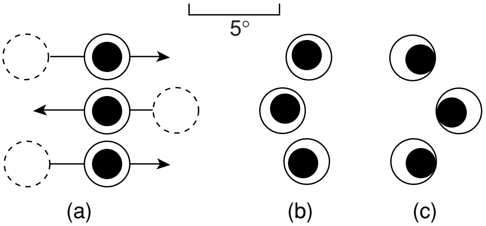


Fig. 24.9 (a) Initial starting positions of three white disks are shown as dashed circles. The disks move horizontally at 11 deg/s, with the top and bottom disks moving to the right and the middle disk moving to the left. As they pass through the position of vertical alignment, three vertically aligned black spots briefly flash up, at the centers of the disks. (b) Artist's impression of how the stimulus looks. The three black dots, which appear to lag behind the centers of the moving disks, appear to lie in a shallow left-pointing V or chevron, whereas the disks themselves appear to lie in a more marked left-pointing chevron. (c) Null settings of the disks and spots at which they look subjectively vertically aligned. (c) is approximately a mirror image of (b).

24.3.1 Three spots show that flash-lag is not caused by differential latencies

One popular explanation of the flash-lag effect attributes it to differential latencies. The idea is that the neural signal from a flashed stimulus is delayed relative to the signal from a moving object (Whitney, Murakami, et al. 2000; Whitney, Cavanagh, et al. 2000). However, I now present some observations that seem to rule out a simple differential latency model. Three white disks move along parallel horizontal paths (Fig. 24.9(a)). The top and bottom disks move to the right and the middle disk moves to the left. At the moment when they pass each other and are vertically aligned, a spot flashes up in the center of each disk so that three vertically aligned dots flash up simultaneously, centered in the three disks. Result: The percept is shown in Fig. 24.9(b). Each spot appears to lag within its disk and looks displaced toward the rear edge of the disk. In addition, the three disks look like a chevron or V pointing to the left, that is, like a snapshot taken of the three disks just *after* they pass each other. The combination of these two effects causes the three spots also to form a shallower subjective chevron that also points to the *left*. In other words, the middle spot looks shifted to the left compared with the top and bottom spots, like a vernier offset (Fig. 24.9(b)). The three moving disks appear to be displaced, as if perceptually registered a little later than their actual occurrence. In addition, the flashed spots are perceptually misaligned with each other, not just with the moving objects. Because there can be no relative delays between the three simultaneously flashed stationary spots, it follows that relative delays cannot account for the subjective vernier misalignment of the spots. Indeed it suggests that this flash-lag effect is not temporal but spatial in nature.

The apparent offsets were measured by a null method. The white disks were each 1.3 deg in diameter and were vertically separated by 1.5 deg between centers. The top and bottom disks started from the left and the middle disk started from the right. The disks moved at 11 deg/sec and covered a horizontal trajectory of 9.5 deg in 860 msec. At the instant when the three disks lined up vertically, a vertical row of three black dots of diameter 0.8 deg was flashed up, each dot lying in the center of one of the disks. A fixation point was provided near this central position (see Fig. 24.9(a)). Observers reported that the disks and black dots appeared to be subjectively offset, with the central disk and the central spot both appearing to be displaced to the left (Fig. 24.9(b)). The observers were provided with four keys. Two of these keys shifted the physical position of the central white disk to the left and right, and two other keys independently shifted the physical position of the central black dot to the left and right. Observers adjusted the positions of the central disk and spot until all three disks and spots appeared to be vertically aligned when the dots were flashed up. They then hit the space bar, which printed out their results. These results were analyzed later offline.

Results are shown in Fig. 24.9, which is drawn to scale. Figure 24.9(a) shows the stimulus setup, in which the top and bottom disks move to the right while the middle disk moves to the left. When the disks are vertically aligned, a black spot is flashed up briefly in the center of each disk. Figure 24.9(b) shows that the spots and the disks appear subjectively to lag behind their true positions, and Fig. 24.9(c) shows the null settings, in which the three disks and spots appeared to be vertically lined up when the central disk was actually displaced to the right by 1.26 deg ± 0.06 deg – almost exactly one disk diameter – and when the central black spot was actually displaced to the right by 0.71 deg ± 0.073 deg – almost one dot diameter (mean ± 1 SE for 4 observers).

The lag of each spot within its disk, and the V-shaped apparent layout of the three disks, comprises the classic flash-lag illusion, and they are consistent with the flashed spots' having a longer latency than the moving disks. One can argue that when a visual snapshot of the three spots is taken and superimposed on the disks, the disks have already moved to a later position. However, the vernier misalignment between the three spots cannot be explained in this way. It cannot be caused by a relative delay, because all spots are flashed up simultaneously. (Incidentally it is unlikely to be an artifact caused by induced movement or position contrast from the moving disks because these hypothetical effects would shift the spots into a subjective chevron that would point to the right, whereas it actually pointed to the left.) We conclude that no relative-latency theory can explain these data.

Cai and Schlag (2001) have made the same argument based on a bar motion experiment. Two bars are traveling in opposite directions and when they line up exactly, a secondary change is added – they either both turn blue or they both get longer, or any number of other things. The apparent location of the changes is offset in the direction of motion (opposite directions). Cai used this new form of illusory conjunction to argue that the flash-lag necessarily has a spatial component because no temporal delay could put the secondary changes where they had never been. His demonstrations can be viewed at http://visionlab.harvard.edu/Members/Rick/Master.htm

Kreegipuu and Allik (2004) also argued that the flash-lag effect has both a temporal and spatial component, and they were able to measure these separately. They presented a moving stimulus that changed its color at a certain moment. The observer indicated, in two different tasks, where and when the color change occurred in relation to a stationary reference flash.

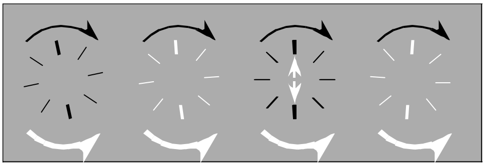


Fig. 24.10 Sixteen radii rotated slowly clockwise (black arrows), changing polarity on every frame. (For clarity, only eight radii are shown.) They seemed to rotate counterclockwise, owing to reversed phi (white arrows). When a radius reached 12 o'clock, a double-headed arrow flashed up (third frame), and observers reported its position. Result: It appeared displaced counterclockwise from the radii, appropriate to the physical clockwise rotation, not to the counterclockwise reversed-phi.

Results show that the perceived time of the color change was not congruent with the perceived location of the color change. The color change was perceived simultaneously with the flash but was shifted in position.

24.3.2 Reversed phi

We also examined the flash-lag effect for a second type of apparently reversed motion, namely reversed phi (Anstis 1970; Anstis & Rogers 1975; Rogers & Anstis 1975).

When two identical pictures are flashed up in sequence, with a small spatial shift between them, apparent motion is seen in the direction of the shift. However, if one picture is the photographic negative of the other, with blacks turned to white and vice versa, the perceived motion is opposite to the actual shift, in the direction of the earlier stimulus. This phenomenon has been called *reversed phi* (Anstis 1970; Anstis & Rogers 1975; Rogers & Anstis 1975). It can be shown (Lu & Sperling 1999) that the motion energy (Adelson & Bergen 1985) in the reversed-phi stimulus is in the direction of the perceived, backward motion.

We prepared a movie in which sixteen radii lay within an annulus with an inner diameter of eccentricity 6 deg and an outer diameter of eccentricity 14 deg (Fig. 24.10). These radii rotated clockwise around their common center at a rate of 3 rev/min (20 sec/rev). On every frame the pattern rotated through 1 deg, and the radii reversed their polarity at a flicker rate of 4.5 Hz, being black on the odd-numbered movie frames and white on the even-numbered frames. Thus, each frame was the photographic negative of its predecessor. We found that the perceived direction of motion depended upon the eccentricity of the radii. If observers tracked a single radius with their eyes, holding the radius in central vision, they correctly perceived it as moving clockwise. However, if they fixated the center of the rotating pattern, the radii, seen eccentrically, appeared to move counterclockwise, owing to reversed phi. (The stimulus was paradoxical in that the motion was perceptually counterclockwise, yet over a long period of fixation it was clear that the locus of the radii was moving clockwise.) Following a period of fixation, if the motion was stopped a strong clockwise motion aftereffect was seen. This was appropriate to the illusory, reversed-phi direction of seen movement, not to the direction of the physical displacement of the radii.

This suggests that the illusory reversed phi was stimulating and adapting neural motion sensors fairly early in the visual pathways (Braddick 1974, 1980).

We combined reversed phi with the flash-lag illusion. Every time a radius reached the 12 o'clock position, which happened sixteen times per revolution, arrows were flashed up at the 12 o'clock and 6 o'clock positions. These were actually aligned with the radii at 12 and 6 o'clock, but any flash-lag illusion would make them look misaligned. Observers were asked to make three qualitative observations in their own time: (1) Was there any apparent misalignment? (2) If so, in which direction was it? (3) In which direction did the radii appear to rotate – clockwise or counterclockwise? Our purpose was to see whether any flash-lag effect was in a direction appropriate to the clockwise physical displacement of the radii or to their perceived (counterclockwise) direction of motion caused by reversed phi.

Five naïve observers were asked to make these three judgments. Results: The flashed arrows appeared to be shifted *counterclockwise* with respect to the moving radii. This perceptual lag was appropriate to the *physical* clockwise rotation of the radii and not to their illusory, reversed-phi counterclockwise rotation. So the flash-lag effect applied to the physical motion, not the illusory reversed motion. We conclude that the flash-lag effect either occurred earlier in the visual system than reversed phi, or else that the flash-lag effect applied to the instantaneous *position* of the radii, regardless of their motion.

24.4 Conclusions

Our results allow us to arrange different visual features into a hierarchical order expressing their strengths in parsing objects.

We have pitted the different moving parts of a chopstick display against each other to see how the visual system parses them, and in particular which parts are perceptually bound together into rigid objects and which are not. These different parts comprise intrinsic and extrinsic line terminators and dots painted on the moving lines. Our results show that these different parts have different strengths – we might say they stick to each other with different strengths of perceptual glue.

Line terminators versus intersections. In the chopstick illusion (Fig. 24.1), the line terminators won out over the X-intersection where the two lines cross each other. The actual motion of the intersection was ignored and replaced by a percept derived from the line terminators. Results were similar for two transparent, overlapping squares in motion. This was not simply because the four terminators outnumbered the one intersection. An array of 22 vertical and 22 horizontal lines contained 484 intersections and only 88 terminators, yet the terminators still won out.

Intrinsic versus extrinsic line terminators. Figures 24.2 and 24.3 showed that intrinsic terminators were very strong cues, extrinsic terminators were very weak.

The flash-lag effect. This was clearly shown to be appropriate to the physical motion of the intersection, not to its perceived direction. It was also appropriate to the physical direction of a reversed-phi stimulus that actually moved in one direction but appeared to move in the opposite direction. These results suggest that the flash-lag occurs early in the visual system, before much motion parsing is done, and is perhaps influenced only by local factors, not by action at a distance, for instance from terminators.

Acknowledgments

Supported by grants from the Academic Senate and Department of Psychology at UCSD. I thank Steve Shimozaki and Dana Ballard for supplying the eye movement data, and Georgina Blanc, Noelle Der-Macleod, Nicole Mead, and Laura Salgado for assistance in collecting the flash-lag data.

References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *J Opt Soc Am A* **2**: 284–299.

Anstis, S. M. (1970). Phi movement as a subtraction process. *Vision Res* **10**: 1411–1430.

Anstis, S. M. (1990). Imperceptible intersections: The chopstick illusion. In A. Blake & T. Troscianko (eds.), *AI and the Eye* (105–117). London: Wiley.

Anstis, S. M. (2003). Levels of motion perception. In L. Harris & M. Jenkin (eds.*), Levels of Perception*. New York: Springer-Verlag.

Anstis, S. M. (2007). The flash-lag effect during illusory chopstick rotation. *Perception* **36**: 1043–1048.

Anstis, S. M., & Rogers, B. J. (1975). Illusory reversal of visual depth and movement during changes of contrast. *Vision Res* **15**: 957–961.

Baldo, M. V., Kihara, A. H., Namba, J., & Klein, S. A. (2002). Evidence for an attentional component of the perceptual misalignment between moving and flashing stimuli. *Perception* **31**: 17–30.

Braddick, O. (1974). A short-range process in apparent motion. *Vision Res* **14**: 519–527.

Braddick, O. J. (1980). Low-level and high-level processes in apparent motion. *Philos Trans R Soc Lond B Biol Sci* **290**: 137–151.

Cai, R., & Schlag, J. (2001). A new form of illusory conjunction between color and shape [Abstract]. *J Vis* **1**(3): 127a, http://journalofvision.org/1/3/127/, doi:10.1167/1.3.127.

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science* **287**: 2036–2038.

Khurana, B., Watanabe, K., & Nijhawan, R. (2000). The role of attention in motion extrapolation: are moving objects "corrected" or flashed objects attentionally delayed? *Perception* **29**: 675–692.

Kowler, E. (ed.) (1990). *Eye Movements and Their Role in Visual and Cognitive Processes*. Amsterdam; New York: Elsevier.

Krauzlis, R. J. (1994). The visual drive for smooth eye movements. In A. T. Smith & R. J. Snowden (eds.), *Visual Detection of Motion* (437–473). London; New York: Academic Press.

Kreegipuu, K., & Allik, J. (2004). Confusion of space and time in the flash-lag effect. *Perception* **33**: 293–306.

Krekelberg, B., & Lappe, M. (2000). A model of the perceived relative position of moving objects based upon a slow averaging process. *Vision Res* **40**: 201–215.

Lisberger, S. G., Morris, E. J., & Tychsen, L. (1987). Visual motion processing and sensory-motor integration for smooth pursuit eye movements. *Ann Rev Neurosci* **10**: 97–129.

Lu, Z. L., & Sperling, G. (1999). Second-order reversed phi. *Perception & Psychophysics* **61**: 1075–1088.

Mackay, D. M. (1961). Interactive processes in visual perception. In W. A. Rosenblith (ed.), *Sensory Communication* (339–355). Cambridge, MA: MIT Press.

Metelli, F. (1974). The perception of transparency. *Scientific American* **230**: 90–98.

Nijhawan, R. (2002). Neural delays, visual motion and the flash-lag effect. *Trends Cogn Sci* **6**: 387–393.

Rogers, B. J., & Anstis, S. M. (1975). Reversed depth from positive and negative stereograms. *Perception* **4**: 193–201.

Shiffrar, M., & Pavel, M. (1991). Percepts of rigid motion within and across apertures. *J Exp Psychol Hum Percept Perform* **17**: 749–761.

Shimojo, S., Silverman, G. H., & Nakayama, K. (1989). Occlusion and the solution to the aperture problem for motion. *Vision Res* **29**: 619–626.

Sinha, P., & Poggio, T. (1996). I think I know that face . . . *Nature* **384**: 404.

Stoner, G. R., & Albright, T. D. (1994). Visual motion integration: A neurophysiological and psychophysical perspective. In A. T. Smith & R. J. Snowden (eds.), *Visual Detection of Motion* (253–290). London: Academic Press.

Ullman, S. (1979). *The Interpretation of Visual Motion*. Cambridge, MA: MIT Press.

Whitney, D., Cavanagh, P., & Murakami, I. (2000). Temporal facilitation for moving stimuli is independent of changes in direction. *Vision Res* **40**: 3829–3839.

Whitney, D., Murakami, I., & Cavanagh, P. (2000). Illusory spatial offset of a flash relative to a moving stimulus is caused by differential latencies for moving and flashed stimuli. *Vision Res* **40**: 137–149.

25

Bridging the gap: a model of common neural mechanisms underlying the Fröhlich effect, the flash-lag effect, and the representational momentum effect

DIRK JANCKE AND WOLFRAM ERLHAGEN

Summary

In recent years, the study and interpretation of mislocalization phenomena observed with moving objects have caused an intense debate about the processing mechanisms underlying the encoding of position. We use a neurophysiologically plausible recurrent network model to explain visual illusions that occur at the start, midposition, and end of motion trajectories known as the Fröhlich, the flash-lag, and the representational momentum effect, respectively. The model implements the idea that trajectories are internally represented by a traveling activity wave in position space, which is essentially shaped by local feedback loops within pools of neurons. We first use experimentally observed trajectory representations in the primary visual cortex of cat to adjust the spatial ranges of lateral interactions in the model. We then show that the readout of the activity profile at adequate points in time during the build-up, midphase, and decay of the wave qualitatively and quantitatively explain the known dependence of the mislocalization errors on stimulus attributes such as contrast and speed. We conclude that cooperative mechanisms within the network may be responsible for the three illusions, with a possible intervention of top-down influences that modulate the efficacy of the lateral interactions.

25.1 Introduction

Localizing an object in the presence of motion is a fundamental ability for many species as a moving object often represents danger or food. In recent years, advances in neurophysiology and psychophysics have substantially increased our understanding of how the visual system calculates the present and future positions of moving objects. New insights have been gained in the last couple of years by analyzing systematic mislocalization errors occurring at different points along the motion trajectory. When observers are asked to localize the initial or final position of a moving stimulus they typically judge the position as shifted forward in the direction of motion. These errors are known as the Fröhlich effect (Fröhlich 1923) and the representational momentum effect (Freyd & Finke 1984; Hubbard & Bharucha 1988), respectively. Another well-established form of mislocalization is the flash-lag effect (Metzger 1932; Nijhawan 1994). When a stimulus is briefly flashed in physical alignment with a continuously visible moving object, observers nonetheless perceive the moving object ahead of the flash. In all three illusions, the position percept does not agree with

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

the physical measurements, indicating that the mechanisms underlying the assignment of an object's location go beyond a mere spatiotemporal filtering of retinal information. Multiple hypotheses pertaining to these illusions have been offered in the past several years. Possible explanations include among others attentional mechanisms (Baldo & Klein 1995; Müsseler & Aschersleben 1998; Kirschfeld & Kammer 1999), active trajectory extrapolation (Nijhawan 1994), differential latencies (Whitney, Murakami, et al. 2000; Maiche et al. 2007), some form of temporal averaging (Eagleman & Sejnowski 2000; Krekelberg & Lappe 2000), or a memory shift of a high-level representation (Hubbard & Bharucha 1988). The lack of agreement on the underlying processing principles may in part be explained by the fact that most of the experimental studies focus exclusively on one illusion. Yet new insights may be gained by elucidating the relationship between the different illusory displacements (Müsseler et al. 2002).

The main purpose of this chapter is to present a network model based on well-known neuronal mechanisms. Within the framework of the model we summarize some existing experimental data and discuss the potential commonalities among the three mislocalization errors. The proposed model consists of a network of excitatory and inhibitory neural populations that encode stimulus position. It implements the fundamental idea that local cortical feedback plays a dominant role in shaping the population representation of a motion trajectory. Internal cooperative mechanisms are in general beneficial for the visual system because they allow, for instance, one to cope with noisy or missing afferent information (Douglas et al. 1995). However, as we shall argue here, they generate in some instances a possible substrate for illusory percepts.

In response to an apparent motion display, the network exhibits a stimulus-locked traveling wave of activity. Lateral interactions mediated by excitatory connections result in a preactivation of neurons encoding future positions. The moving object is thus processed more efficiently compared to a flashed object as in a flash-lag display, resulting in a differential processing delay. We have recently reported neural trajectory representation in the primary visual cortex of cat (Jancke et al. 2004b) that showed such a path-dependent facilitation. If the cooperative mechanisms within the network are sufficiently strong, the dynamic transformations sustain for some time upon stimulus offset (Erlhagen & Jancke 2004). As a result, neurons encoding positions displaced forward in the direction of motion become active. The population response thus represents a possible neural substrate for a stimulus position that is perceived but not sensed directly.

To test the idea that recurrent interactions constitute a common low-level mechanism for explaining the illusions, we proceed as follows. We first use the experimentally observed traveling waves in cat primary visual cortex to calibrate the model parameters that represent the spatial ranges of the lateral interactions. Our working hypothesis for discussing the three mislocalization phenomena is that spatiotemporal characteristics of the build-up, middle, and decay phase of the wave are related to the Fröhlich effect, the flash-lag effect, and the representational momentum effect, respectively. We test this hypothesis by directly comparing model predictions and experimental findings when stimulus attributes such as contrast and speed are systematically varied. There is experimental evidence that changes in

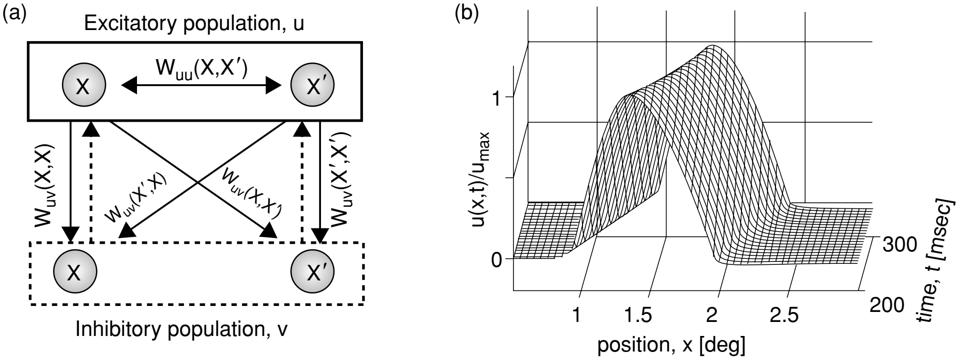


Fig. 25.1 (a) Schematic representation of the connectivity between any two nodes x and x' of the model network consisting of an excitatory (u) and an inhibitory (v) subpopulation of neurons coding for position. For details see the text. (b) Traveling wave in response to a stimulus of width $2\sigma_S = 0.4$ deg and intensity $A_S = 13.2$ moving with the apparent speed $v = \Delta x/\Delta t = 40\,\text{deg}/s$ (spatial displacement $\Delta x = 0.4$ deg, frame duration $\Delta t = 10$ msec). Model parameters were: $\sigma_u = 0.3$ deg, $\sigma_v = 0.4$ deg, $A_u = 4.65$, $A_v = 3.99$, $\beta = 1$, $u_f = u_g = 0$, $h = -3$. To adjust the spatial scale in the model to the experimental units we have chosen 10 pixel = 0.2 deg.

the task demands may alter the magnitude of mislocalization errors (Müsseler et al. 2002; see also the discussion in Kreegipuu & Allik 2003), suggesting that the effects cannot be explained by a feedforward architecture alone. We propose and test a neuroplausible mechanism that allows altering the efficacy of the recurrent interactions based on top-down influences.

25.2 The dynamic model

25.2.1 Model architecture

There are two main hypotheses about the neural mechanisms underlying the processing of stimulus position that have guided our modeling work (for a detailed discussion see Erlhagen et al. 1999; Jancke et al. 1999). First, information about stimulus position is encoded in visual brain areas by the distributed activity pattern of large neural populations, rather than by single neurons. Each cell of a population is tuned to a specific position in visual space, and its level of activity defines the extent to which the information is present. A second hypothesis concerns the role of cortical interactions in shaping the response properties of neural populations. The fact that the largest input to cortical cells comes from neighboring cells rather than from feedforward afferents suggests that massive excitatory feedback counterbalanced by cortical inhibition plays a central role for the processing of stimulus attributes such as position (Douglas et al. 1995).

In the model, neurons split into an excitatory and an inhibitory subpopulation. They are organized as layers of neurons densely covering the stimulus dimension. Because we focus on experiments with horizontal motion displays, the model is one-dimensional.

The connection schema for any two neurons of the network is sketched in Fig. 25.1(a). An excitatory neuron tuned to position x integrates activity from neighboring excitatory neurons via lateral connections and projects to neurons of the inhibitory pool with similar and dissimilar receptive field centers x and x', respectively. This model architecture is in line with anatomical and physiological work demonstrating that direct reciprocal connections exist primarily between pairs of excitatory pyramidal cells. Axon collaterals of these cells, on the other hand, may also target inhibitory interneurons (Gilbert 1995). Each inhibitory neuron x spatially summates incoming activation from excitatory neurons. For simplicity, we assume that it feeds its activation back to the excitatory pool only locally, that is, to a neuron with the same receptive field center x. Note that the implementation of spatially extended inhibitory feedback would not change the qualitative conclusions of our work.

Like many other models of cortical function (for an overview see Dayan & Abbott 2001), cortical interactions is assumed to depend on the functional distance of the cells and defined by the feature coded by them. The model parameters describing the recurrent interactions are adjusted such that neurons with similar receptive field centers excite each other, whereas inhibition dominates for larger distances. This interaction pattern, known as "Mexican-hat" organization, guarantees a sharply tuned excitation profile as a network response to an afferent input that carries information about the visual location of a stimulus.

Assuming that the number of excitatory and inhibitory neurons is large and that their receptive field centers densely cover the visual field, the mean activity at time t of an excitatory neuron and an inhibitory neuron tuned to horizontal position x can be described by two continuous functions $u(x, t)$ and $v(x, t)$, respectively. To model the dynamics of the neural population, we use the model class of neural fields first introduced and analyzed by Wilson and Cowan (1973). Neural field models are system-level models adequate to describe the mean activity of large populations of neurons without referring to a detailed level of physiological realism. The following differential equations (Jancke et al. 1999; Erlhagen & Jancke 2004) govern the evolution of the activation variables u and v:

$$\begin{aligned}
\tau \frac{d}{dt} u(x,t) &= -u(x,t) + h + S(x,t) + g(u(x,t)) \\
&\quad \times \left[\int w_{uu}(x - x') f(u(x',t)) dx' - v(x,t) \right] \\
\tau \frac{d}{dt} v(x,t) &= -v(x,t) + \int w_{uv}(x - x') f(u(x',t)) dx'
\end{aligned}$$

where du/dt and dv/dt represent the changes in mean activity over time. The parameter $\tau > 0$ is used to adjust the time scale of the field dynamics to the experimentally observed time scale. The afferent input $S(x, t)$ to the excitatory population is modeled as a Gaussian profile. Its space constant, σ_s, reflects the half-width of the stimulation, whereas the amplitude, A_S, is assumed to change as a function of the luminance contrast of the external stimulus. The constant h defines the resting level to which the population activity relaxes without external input. The integral terms describe the spatial summation of excitation in the two layers. The spatial interactions fall off with increasing distance between field sites:

$$w_{ui}(x - x') = A_i \exp\left(-\frac{(x - x')^2}{2\sigma_i^2} \right), (i = u,v),$$

where the choice of the relative amplitudes and spatial ranges, $A_u > A_v$ and $\sigma_u < \sigma_v$, implements the Mexican-hat pattern. Only sufficiently activated neurons contribute to the interaction. The nonlinear activation function $f(u)$, which gives the mean firing rate for a given level of activation, is taken as a monotonic function of sigmoid shape going from 0 to 1:

$$f(u) = \frac{1}{1 + \exp(-\beta(u - u_f))}.$$

The parameter u_f determines the position of the maximum slope of the function f, and β controls the value of the maximum slope.

The field dynamics exhibit a threshold behavior. Starting from the stable resting state h, only a sufficiently strong afferent input is able to drive excitation to a level that triggers the self-stabilizing forces within the network. The activation in the excitatory layer is normalized relative to the threshold excitation level $u_{TH} = 0$. The resting state is thus chosen to be negative ($h < 0$), and consequently negative u-values describe subthreshold activity.

An increasing body of experimental evidence suggests that cognitive factors like attention or task demands can alter the efficacy of the lateral connectivity in primary visual brain areas (Li et al. 2004). There is certainly a need for neural mechanisms that allow for flexible visual processing without referring to learning-based synaptic reorganization that is believed to take place on a longer time scale. The neural sources for such top-down influences are not known. A number of architectures have been proposed that, in principle, enable changing the functional properties of recurrent networks in an efficient way (Hahnloser et al. 1999). We have implemented a simple shunting mechanism (for an overview see Grossberg 1988) that can alter the gain of the network response to an afferent stimulus. The recurrent interaction of the u-layer is gated by a nonlinear function $g(u)$ that is also of sigmoid type with parameters u_g and β. We simulate a top-down influence by gradually changing the position of the maximum slope u_g. This allows us to control in a flexible manner the spatial spread of the suprathreshold population response and thus the extrapolation properties of the network (Erlhagen 2003).

25.2.2 Choice of model parameters

We study how the neuronal population response, which is shaped by the cooperative mechanisms within the network, interacts with an external input representing a stimulus in motion. When a brief, localized input of adequate intensity is applied, the network develops a localized activity pattern in position space known as an “active transient” (Wilson & Cowan 1973). Due to the recurrent excitation, the activity continues to grow in amplitude and width upon stimulus offset. It reaches a peak value and then decays back to resting level driven by the increasing feedback inhibition. If a flashed stimulus is part of a motion paradigm, the population representation of that stimulus will interact through the lateral connections with representations of preceding and succeeding stimulus frames. As shown in Fig. 25.1(b), the delicate interplay between the excitatory and inhibitory feedback loops may result in an activity wave that propagates with the velocity of the inducing stimulus in the direction of motion. This locking to the stimulus occurs for a whole range of velocities (Ben-Yishai et al. 1997; Giese & Xie 2002), thus permitting the study of the effect of velocity changes on the spatiotemporal characteristics of the wave within a single population. The velocity range is determined by the interaction structure of the network. It covers the speed of a spontaneous wave that evolves in the absence of a time-dependent input for a sufficiently low threshold u_g of the gating mechanism (Erlhagen 2003). Beyond the characteristic range, the wave loses stability and a population response with a quasi-periodic amplitude modulation emerges, suggesting that multiple neural networks with different speed selectivity might exist.

We use the neural trajectory representations, which we have recently described in the primary visual cortex of cat, to adjust the model parameters that define the interaction processes. The velocity range tested in the experiments (4–40 deg/sec) constrains the spatial ranges for excitation and inhibition, $\sigma_u < \sigma_v$. The amplitude ratio $A_u/A_v > 1$ is adjusted to reproduce the mean amplitude and width of the neural waves. It is worth mentioning that the whole set of model parameters used for the present study is in agreement with our previous modeling of neural interaction effects probed with stationary displays (Jancke et al. 1999).

The link to the perceptual mislocalization errors is made by taking the peak position of the traveling wave as an estimate of stimulus position. Our working hypothesis is that the wave model fitted to reproduce the neural data recorded in the visual cortex will allow us to test whether such a self-stabilized representation may qualitatively explain the three visual illusions. To also reproduce the magnitude of the mislocalization errors observed in the psychophysical experiments we make two adjustments to the model parameters. A slightly larger amplitude ratio $A_u/A_v > 1$ is used. This favors the recurrent excitation over the inhibition and results in a stronger facilitation effect along the motion trajectory. A second adjustment concerns the time scale of the field dynamics. It is fixed to $\tau = 15$ msec for the population data, which is in the range of the membrane time constant of a neuron (Abeles 1991). To cover also a perceptually relevant time scale we adapt the value to $\tau = 35$ msec for the modeling of the localization errors. This ensures that the duration of the active transient response to a brief afferent input reflects the persistence of a flashed stimulus in the visual system (100–150 msec, Coltheart 1980). The identical set of field parameters is used for the simulation of the three visual illusions. Only the gating threshold u_g is modified in some simulations to model assumed top-down influences on the efficacy of the recurrent interactions. The stimulus dimension and the spatiotemporal properties of the linear motion displays (frame duration Δt, spatial displacement Δx) are adapted to meet in each case, as close as possible, the conditions of the psychophysical experiments.

25.3 Modeling results

25.3.1 Motion trajectories in primary visual cortex

To address how motion trajectories are represented at the level of primary visual cortex we recently used a population approach that pools spiking activity of many neurons in cat area 17 (Jancke et al. 2004b). Following our earlier study with stationary displays, our working hypothesis was that the observed nonlinear interaction effects between neuronal representations of adjacent stimuli should manifest also when the afferent input is time-dependent. In those experiments we used small squares of light (0.4 deg) that moved with different velocities ranging from 4 to 40 deg/sec along a horizontal line. We employed an optimal linear estimator technique that allows reconstructing with high spatial and temporal resolution the stimulus position from a pool of broadly tuned neurons (Salinas & Abbott 1994; Erlhagen et al. 1999). The general idea behind the construction of population distributions is that the activity of each cell in each trial is treated as a vote for its tuning curve (or in case of an optimal estimator a so-called basis function obtained by a template-matching procedure). The summation of all votes weighted by the firing rate gives the population response in parametric space. Despite the fact that adjacent stimuli activate highly overlapping cell populations, the reconstructed motion trajectories reveal that the peak of the population response represents well even small changes in stimulus position.

Two examples of neuronal trajectory representations that differ in velocity and direction are shown in a space–time diagram in Fig. 25.2(a). One striking characteristic of these traveling activity waves is that the mean activation level increases with stimulus velocity. As shown in Fig. 25.2(b), the network model with adequately adjusted spatial ranges of lateral excitation and inhibition can explain this finding. In the model, the spread of excitation is followed by a wave of inhibition that reduces the amplitude as well as the duration of the population response. This suppressive effect increases with lower speed. Moreover, the model predicts, in line with the experimental findings, that the response to the stimulus train with the highest tested velocity reaches approximately the mean activation level of the response to the stimulus flashed in isolation. A second characteristic of the neuronal trajectory representations captured by the dynamic model is that the localized activity profile locks to the stimulus with a speed-dependent spatial lag. To allow for a direct comparison of modeling and experimental results (Fig. 25.2(c)), a constant time interval of 25 msec was added before stimulus onset in all simulations. This time window represents the average temporal delay between the stimulus presentation and the onset of the population response in primary visual cortex (Jancke et al. 1999). There are three findings of particular interest for the discussion of the localization errors. First, a near compensation of processing delays for the slowest speed can be observed. Second, for the range tested the speed dependence of the spatial lag is roughly linear. However, the most important finding is that the peak latency of the population response to the stimulus in motion was shorter by about 16 msec compared to the response when the stimulus was flashed in isolation (Jancke et al. 2004b). The observed differential latency in primary visual cortex gives direct physiological support for explanations of visual illusions that stress the importance of the time at which an object is perceived (for a review, see Whitney 2002). It is important to note that this latency reduction is observed even for the stimulus train with highest velocity. Here, the individual stimulus frames do not significantly overlap. This excludes a simple explanation of the differential latency based on the fact that the spatially extended moving stimulus may start to trigger retinal cells earlier than the flash.

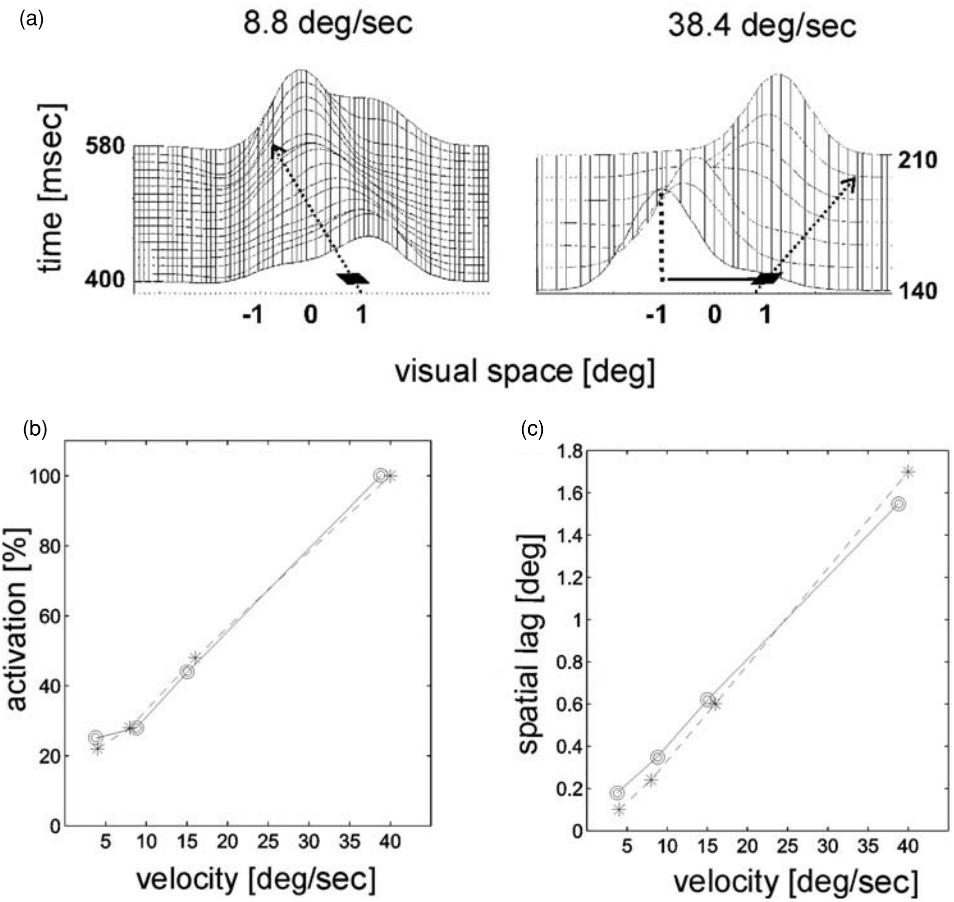


Fig. 25.2 (a) Two examples of cortical trajectory representations: a square stimulus of width 0.4 deg moved with speed 8.8 deg/sec (left) and 38.4 deg/sec (right). The reconstructions of the traveling waves are based on the activity of 178 neurons recorded in the central visual field representation of cat area 17. (b) Comparison of model predictions and experimental findings: dependence of the mean activity level of the wave on stimulus speed. The value averaged over both directions tested in the experiments is shown for the model simulations (asterisk) and the neuronal reconstructions (circle). Different movement velocities were induced by using identical temporal presentation rates $\Delta t = 5$ msec and adequately adapted spatial displacements. Model parameters were like in Fig. 25.1 except $A_u = 4.52$, $A_v = 4.39$, $u_f < u_g = 3$. (c) Dependence of the spatial lag on stimulus speed. The mean spatial displacement between the current stimulus position and the peak position of the traveling wave is plotted as a function of speed for the model simulations (asterisks) and the neuronal representations (circle).

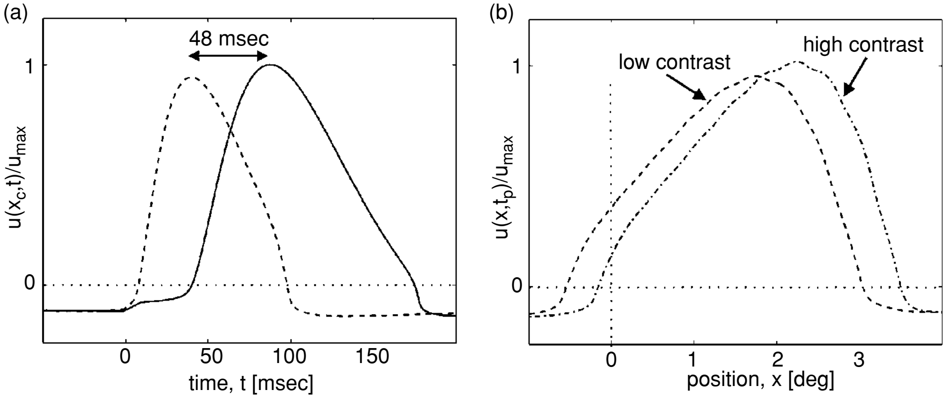


Fig. 25.3 (a) Comparison of the network response to a flashed stimulus ($\Delta t = 10$ msec, $\sigma_S = 0.2$ deg, $A_S = 6.6$) when it is presented at time $t = 0$ in isolation (solid line) or as part of a motion display ($\Delta x = 0.4$ deg, $\Delta t = 10$ msec, dashed line). The time course of the maximal exited neuron at the center position x_c is shown. The activity level $u_{TH} = 0$ indicates the threshold for triggering the active transient response. The same set of model parameters as in Fig. 25.1 was used. (b) Dependence of the flash-lag illusion on the contrast of the moving stimulus. The position of the traveling wave at the time t_p of the peak response to the stationary flash is compared with the wave position when the stimulus intensity A_s of the moving stimulus was increased by a factor of 2, $A_S = 13.2$. The wave appears further ahead of the flash position $x_c = 0$ for the "high contrast" (HC) compared to the "low contrast" (LC) stimulus.

25.3.2 The flash-lag effect

The flash-lag effect (FLE) describes a visual illusion wherein a moving object is perceived as being ahead of a stationary flashed object when the two retinal images are physically aligned. It was first discovered some 80 years ago and has often been explained as being due to differential perceptual latencies for the flashed and the moving stimulus (Metzger 1932; Purushothaman et al. 1998; Whitney, Murakami, et al. 2000; Maiche et al. 2007). If the time to perception for the flash were longer, the moving object should appear ahead of the flash position. Taking the peak latency of the flash response as a time maker, the findings of our population study suggest a neural correlate for this explanation because the peak of the traveling wave has already passed the flash position. However, the observed latency difference of 16 msec is not sufficient to explain the full range of flash-lag effects for the majority of experimental studies, which find differences in the range of 45–80 msec (Krekelberg & Lappe 2001). We therefore explored in the model simulations of the impact of a larger amplitude ratio $A_u/A_v > 1$ of the excitation relative to the inhibition on the path-dependent facilitation. An important constraint for the modeling comes from the finding that no significant mislocalization further ahead in motion direction occurs when with the disappearance of the flash the moving object also disappears (Whitney, Murakami, et al. 2000; but see Fu et al. 2001). This means that the cooperative mechanisms within the network should not be strong enough to sustain the wave beyond the vanishing position. In Fig. 25.3(a) we compare the time course of the maximum excited neuron in response to the stimulus flashed in isolation (solid line) and when it is part of a linear motion paradigm (dashed line). Stimulus onset is for both cases t = 0 msec. Due to the spread of subthreshold activity from preceding frames, the stimulus in motion triggers immediately the suprathreshold response, whereas it takes much longer when the stimulus is flashed without spatiotemporal context. Because the network inhibition following the excitation also starts earlier, the peak latency appears to be reduced by 48 msec in this example. This value is in agreement with the 45 msec inferred by Whitney and colleagues from the findings of their linear motion paradigm. Consistent with the differential latency hypothesis, the flash-lag effect depends systematically on the luminance of the object (Purushothaman et al. 1998). The visual latency for a given stimulus is believed to vary inversely with luminance (Lennie 1981). Consequently, the flash-lag effect is predicted to increase if the luminance of the moving object but not that of the flash is increased. Reducing instead the flash latency by applying a higher flash contrast decreases this lead. The network model captures the latency dependence on contrast because stronger afferent inputs reach the threshold u_{TH} for the self-stabilized population response earlier. As depicted in Fig. 25.3(b), the traveling wave model is qualitatively in line with these experimental observations concerning the flash-lag effect. The snapshot of the wave at the peak time of the flash response is further ahead of the flash position for the stimulus with higher intensity A_s.

A latency advantage of about 50 msec represents the maximal value that can be achieved with the present parameter settings. It is important to stress that this advantage is the result of cooperative mechanisms within a single stage of cortical visual processing. Taking the afferent pathway from retina to cortex into account may allow explaining even larger flash-lag effects. Mechanisms like contrast gain control in populations of retinal cells (Berry et al. 1999) or a structured summation over excitatory feedforward input from hierarchically lower processing stages (Baldo & Caticha 2005) have been proposed as explanations for the advance of the moving stimulus over the flash. Although the largest input to cortical cells comes from neighboring cells at the same stage of processing, afferent input may contribute to the preactivation of neurons encoding future stimulus positions, and thus to the path-dependent facilitation. The observation that the flash-lag may turn into a flash-lead for high-contrast flashes (Purushothaman et al. 1998) requires the assumption of differential delays from retina to cortex (Baldo & Caticha 2005). In the present simulations the intensity parameter, A_S, was varied but not the onset of the afferent input, $S(x, t)$, to the excitatory population.

A second stimulus parameter that has been systematically analyzed in experimental studies is stimulus speed. A linear increase of the flash-lag illusion with speed has been reported (Nijhawan 1994; Whitney, Murakami, et al. 2000; for a review see Krekelberg & Lappe 2001). The wave model provides a natural explanation for this finding because the time interval that elapses before the population peak reaches stimulus position is approximately constant for all velocities (compare Fig. 25.2(c)).

25.3.3 The Fröhlich effect

Fröhlich (1923) was the first who studied systematically the phenomenon that a slit of light moving on a track is not seen immediately after it emerges from behind a screen

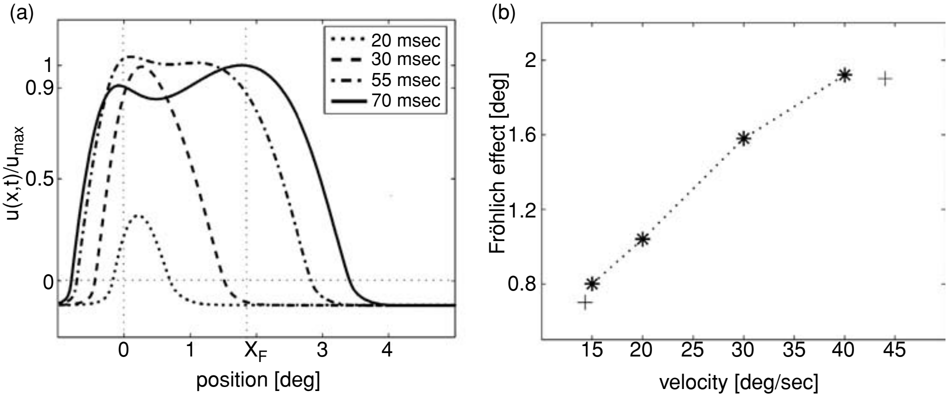


Fig. 25.4 (a) Four snapshots of the buildup phase of the wave used in Fig. 25.3 are plotted. The position x_F indicating the start of a coherently propagating activity pattern is taken as a correlate for the Fröhlich effect. (b) Comparison of model predictions (asterisk) and experimental findings (plus) of the Fröhlich effect as a function of speed. The experimental data were estimated from Fig. 2 in Müsseler & Aschersleben (1998). Model parameters were like in Figs. 25.1 and 25.3, stimulus parameters were $\sigma_S = 0.25$ deg, $A_S = 13.2$. Different movement velocities were induced by using a fixed temporal presentations rate $\Delta t = 3$ msec and adequately adapted spatial displacements.

but only after passing through a certain distance. Recently, Müsseler and Aschersleben (1998) established the existence of this illusory displacement of the onset position with a computer-generated motion paradigm. Early temporal accounts to explain this effect were based on the idea that during the time it takes to perceive the object ("sensation time") it has already moved a certain distance (for a detailed discussion see Müsseler & Aschersleben 1998). In the network model, the build-up of the traveling wave is not immediate and occurs at positions subsequent to the object's initial position. The straightforward idea behind the modeling work is thus to qualitatively and quantitatively discuss the when and where of the wave evolution in relation to the Fröhlich effect.

Figure 25.4(a) shows four snapshots describing the build-up of the population response to the motion display used to discuss the flash-lag illusion (see Fig. 25.3). The activity pattern first crosses the threshold for the self-sustaining mechanisms at a position that is shifted forward in motion direction relative to the starting position $x = 0$. Due to the local excitatory loops within the network, the activity continues to grow both in amplitude and width. A maximum activation level is reached when the local inhibition starts to dominate the processing. Subsequently, an activity peak evolves that starts to follow the stimulus with the respective speed. Note that it still takes about 30 msec until the balance of excitation and inhibition stabilizes the final shape of the traveling wave. It is important to stress that only at the point in time of the peak formation the spatiotemporal activity pattern carries in a reliable manner the information about a stimulus in motion. As our modeling and neurophysiological data show (Jancke & Erlhagen, unpublished), the build-up within the first 70 msec time window resembles the population response to a briefly flashed, spatially extended bar.

To directly compare model predictions with experimental findings, we read out the peak position at the time when the "tail" of the activity profile has decayed to about 90% of its maximal activation. The dependence of this reference value x_F on stimulus parameters such as contrast and speed qualitatively reflects the experimental observations. Fröhlich (1923, p. 73) pointed out that increasing stimulus luminance reduces the illusory displacement at movement onset. Consistently, position x_F comes closer to the starting position with increasing intensity A_s, indicating an earlier onset of the trajectory representation. We have adjusted this parameter to also quantitatively reproduce the dependence of the Fröhlich effect on speed. In the horizontal motion paradigm of Müsseler and Aschersleben (1998), the stimulus was moved at two different velocities, 14.3 deg/sec and 44 deg/sec, in the two directions. In Fig. 25.4(b) we compare the magnitude of the experimentally observed displacement averaged over both directions (plus) with model predictions for four velocities in the range between 14.3 deg/sec and 44 deg/sec (asterisk). The modeling results reproduce well the observed increase of the effect with speed.

A robust mislocalization in motion direction has been also reported when the first frame of the motion display appears simultaneously with a flash. Because in some studies of this so-called flash-initiated cycle (Khurana & Nijhawan 1995) the shift in onset position appeared to be comparable in magnitude to that of the standard flash-lag illusion, it has been argued that the same processing mechanisms might cause the two localization errors (Khurana & Nijhawan 1995; Eagleman & Sejnowski 2000; see also Nijhawan et al. 2004). The differential latency account would again suggest the flash as a time marker for reading out the population activity in response to the moving stimulus. However, Eagleman and Sejnowski (2000) convincingly showed that presenting the flash 50 msec before the onset of the motion does not affect the error. The authors propose instead that the flash "resets" motion integration. Consequently, like at motion onset an internal position representation (e.g., a traveling wave) has first to be built up. On this view, the flash-lag illusion is just a variant of the Fröhlich effect. Other studies, however, that also investigated directly the relation of the two localization errors (albeit using higher speeds) reported a significant difference in magnitudes (Müsseler et al. 2002; Kreegipuu & Allik 2003), suggesting different underlying mechanisms. The network model makes a clear prediction to decide this open question. With increasing stimulus contrast, the traveling wave starts earlier in space and time. Simultaneously, the wave appears further ahead of the population representation of a physically aligned flash (compare Fig. 25.3(b)). The model thus predicts the opposite effects of changes in stimulus contrast on the Fröhlich and flash-lag illusions.

25.3.4 The representational momentum effect

When observers are asked to remember the final position of an object in motion they typically misremember it as further along the implied trajectory. Analogous to Newton's first law of motion, this form of motion extrapolation has been termed "representational momentum" (Freyd & Finke 1984). The momentum metaphor refers to the notion that like a physical object the dynamic internal representation of position cannot be halted immediately

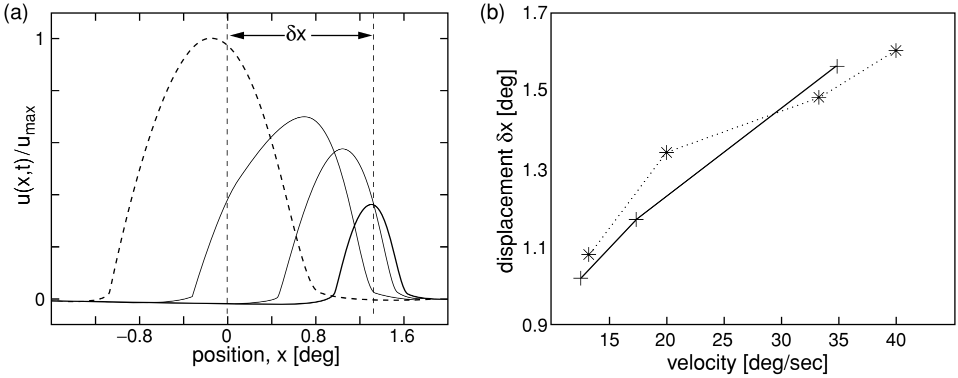


Fig. 25.5 (a) Extrapolation to positions that were not physically stimulated. Four snapshots of the wave at the end of the motion trajectory are shown. The leftmost profile (dashed line) represents the wave at the time of stimulus offset. The forward displacement δx is defined as the spatial distance between the actual vanishing position $x = 0$ and the peak position of the wave when it stops to travel (rightmost profile). Model parameters were like in Figs. 25.1 and 25.3 except a lower gating threshold $u_f = 0 > u_g = -0.25$. (b) Comparison of model predictions (asterisk) and experimental findings (plus) of the representational momentum as a function of speed. The experimental data was estimated from Fig. 1 in Hubbard and Bharucha (1988) and Fig. 1 in Hubbard (1990). Model parameters were like in Figs. 25.1 and 25.3, stimulus parameters were $\sigma_S = 0.45$ deg, $A_S = 10$. Different movement velocities were induced by using a fixed frame duration $\Delta t = 3$ msec and adequately adapted spatial displacements.

upon stimulus offset. The effect has been originally found in implied motion paradigms but has been later replicated with continuous motion displays (Hubbard & Bharucha 1988; Hubbard 1990).

Depending on the efficacy of the lateral interactions, the traveling wave may overshoot the vanishing position. The population response thus suggests the presence of a moving object at positions that were never physically occupied. For the simulation example shown in Fig. 25.5 we used the identical set of model parameters as for the simulation of the flash-lag illusion (Fig. 25.3) but reduced the gating threshold u_g for the lateral interactions. As depicted by the snapshots in Fig. 25.5(a), the wave still lags behind stimulus position at the time of stimulus cessation (dashed line). The spatial lag is, however, smaller compared to the simulation shown in Fig. 25.3, giving further support for the notion that lateral interactions are an efficient means to compensate for processing delays. Upon stimulus offset, the wave loses speed and amplitude. Finally, the activity profile stops to travel and decays back to resting level. We use the peak position representing the offset of the continuous propagation as a correlate for the representational momentum. The forward displacement $\delta x > 0$ is defined as the difference between that peak position and the actual vanishing position.

In Fig. 25.5(b) we compare model predictions with the experimental findings reported by Hubbard and Bharucha (1988) and Hubbard (1990) using a horizontal motion display. The experimental data (plus) for the three speeds, 12.5, 17.4 and 34.8 deg/sec, represent the average value of the observed forward displacement for the four tested subjects. The modeling results (asterisk) for speeds within this range reproduce well the order of magnitude of the effect and the increase of the forward displacement with speed. Note that for the set of parameters used, the range of stimulus velocities that leads to a smoothly traveling wave is between 10 and 40 deg/sec. Beyond this range, the amplitude oscillations of the population response make the overshoot dependent on the exact vanishing position.

There is evidence that stimulus attributes such as contrast or shape (e.g., blurred edges) may affect the overshoot of the vanishing position of a moving target (Fu et al. 2001). Also the population response of the model network shows this dependence on stimulus parameters. A stronger afferent stimulation, for instance, leads to a larger forward displacement. However, the increase is rather small ($< \sigma_s$) and cannot explain alone the order of magnitude of the momentum effect.

In the model simulations, a change in efficacy of the lateral interactions by adapting the gating threshold allows us to explain the seemingly conflicting psychophysical data about position judgment at the vanishing position. Under the condition of the flash-lag paradigm no overshooting is observed. The only, but important, difference to the representational momentum paradigm is the judgment relative to an accompanying flash. In a study designed to directly compare the flash-lag illusion and the representational momentum effect, Müsseler and colleagues (Müsseler et al. 2002) found evidence for some kind of cognitive control over the position judgment. They showed that a verbal manipulation of the task relevance of the accompanying flash (“ignore the flash”) may gradually change the pattern of mislocalization at the end of the motion trajectory. When trying to compare the position of the moving stimulus with an unpredictable position in space, the temporal facilitation of subsequent stimuli becomes less important because no latency compensation is needed for this task. One may hypothesize that the instruction cue about the type of position judgment (relative or absolute) generates feedback to lower visual areas, thereby altering the efficacy of the lateral interaction loops before the stimulus is processed (Lamme & Roelfsema 2000).

In this context it is worth mentioning that in Hubbard and Bharucha (1988) all subjects showed the increase of the forward displacement with speed. However, there is a large difference in magnitude across subjects. Assuming that many individual top-down influences (e.g., attention, experience, or context) can alter the efficacy of the lateral interactions (Li et al. 2004) may explain this finding.

25.4 General discussion

We have suggested that the spatiotemporal dynamics of a network model that incorporates plausible assumptions about the local cortical connectivity accounts for perceptual correlates of localization errors observed in various motion paradigms. In its functional architecture the network model reflects converging lines of physiological and anatomical evidence that single neurons are not passive filters but dynamic entities with response properties depending on the collective behavior of large populations of cortical cells (Fitzpatrick

2000). The interplay between afferent inputs and local feedback loops results in a wave of activation that locks to the moving stimulus. The cooperative mechanisms within the network explain the critical dependence of the visual illusions on the stimulus attributes of contrast and velocity.

25.4.1 Facilitation through preactivation

The fundamental finding of our population study in the primary cortex was that the neural trajectory representations reveal a reduced latency when compared to a flashed stimulus. The threshold mechanism that causes such a temporal facilitation in the model network relies on the assumption that spreading subthreshold activation leads to a preactivation of surrounding neuronal populations (Kirschfeld & Kammer 1999). Consequently, the threshold for spike generation is reached faster compared to a population starting the processing of the afferent input from resting level. There is a growing body of experimental evidence from optical imaging and intracellular recording studies that supports the existence of a subthreshold depolarization wave temporally ahead of the afferent input (Grinvald et al. 1994; Bringuier et al. 1999; Jancke et al. 2004a).

One important question discussed in the context of the differential latency account to the flash-lag illusion is whether the temporal facilitation is omnidirectional or restricted to the specific motion trajectory (Nijhawan et al. 2004). Whitney and colleagues (Whitney, Murakami, et al. 2000; Whitney, Cavanagh, et al. 2000) found no significant difference in the effect even during unpredictable changes in direction including motion reversal. Because the interaction structure of the model network is completely symmetric there is no a priori preference for a certain direction. However, an asymmetry is introduced by the fact that the propagating excitation is followed in time and space by an inhibition wave. For the particular case of motion reversal this means that the afferent stimulus interacts with local inhibition caused by the previous stimulation of the same position in visual space. As a result, a coherently traveling wave has first to build up, suggesting that the observed flash-lag may be caused by the same mechanisms as for motion initiation and not by differential latencies. However, for a two-dimensional extension of the model network we expect that the circular spread of excitation mediated by an omnidirectional interaction structure may be sufficient to guarantee a temporal facilitation effect for a whole range of direction changes (excluding the particular case of motion reversal).

25.4.2 Readout of positional information

We have suggested that the traveling wave constitutes a common neural mechanism that relates the three illusory displacements observed with objects in motion (see also Müsseler et al. 2002). The localization at motion initiation and motion offset is defined by the starting and stopping position of the wave, respectively. The Fröhlich effect is explained by the fact that the starting position appears to be shifted forward with respect to the first physical position of the stimulus. The representational momentum effect has its correlate in the overshooting of the last stimulus position. Finally, the differential latency of the flash and the moving stimulus determines the localization error at a midposition of the trajectory. The "reading out" of the neuronal activity pattern is done from a snapshot for simplicity. It is important to note that an averaging of positions over a whole time interval (Eagleman & Sejnowski 2000) could have been chosen as well without qualitatively changing the overall pattern of results. For instance, using the time window defined by the suprathreshold flash response would explain a larger flash-lag compared to a readout at the time of peak latency. This is due to the systematic asymmetry of the temporal evolution with respect to the peak.

There is no reason to believe that the three position estimates will define visual illusions of the same order of magnitude. In the only study thus far that directly compared the three illusory displacements in a single experimental setting, Müsseler and colleagues (Müsseler et al. 2002) found different sizes for the Fröhlich, the flash-lag, and the representational momentum effect. Inspired by our work, the authors also introduced the notion of a wave as a common mechanism that might explain their data.

25.4.3 Motion extrapolation

In the last 10 years, the flash-lag effect has been the object of an intense debate. Nijhawan (1994) renewed interest in this effect by explaining it as an attempt by the visual system to compensate for processing delays. These delays cause the neuronal representation to spatially lag behind the actual position of the moving object. The extrapolation hypothesis states that the visual system compensates for this spatial lag by actively shifting the coded position forward along the trajectory of motion. Consistent with this hypothesis, Berry and colleagues (1999) demonstrated convincingly that the peak of the population response of retinal ganglion cells may lead the stimulus position in visual space. However, they also reported a decrease of this spatial lead with increasing speed of the moving object. The findings of our population study in primary visual cortex suggest that the spatial lead may change into a spatial lag for sufficiently high speeds. Although this dependence on speed does not support the mechanism originally proposed by Nijhawan, the model simulations show that a mechanism based on lateral interactions may explain a substantial compensation of processing delays.

A task-related alteration of the efficacy of the lateral interactions by means of the proposed shunting mechanism allows controlling the extrapolation properties of the network. We propose such a top-down contribution as an explanation for the lack of effect in the flash-terminated cycle. In the flash-lag paradigm the cooperative forces within the network are tuned to prevent the wave from overshooting the vanishing position. For tasks that require the accurate localization of a moving object, the lateral interactions can be tuned to bring the peak of the neural trajectory representation close to the actual stimulus position. However, the system pays the price that this predictive representation overshoots the final position of the object when it abruptly vanishes (compare Fig. 25.5(a)).

Changes in the efficacy of the lateral interactions affect not only the decay but also the buildup of the wave. Consequently, the network model predicts a correlation between the error at the start and the end of the motion trajectory. This correlation has indeed been found in studies in which the subjects were asked to indicate both the perceived onset and offset position of the moving stimulus (Hubbard & Motes 2002; Müsseler et al. 2002; Thornton 2002). The fundamental finding was that the perceptual overshooting at the end of the trajectory is accompanied by the disappearance of the Fröhlich effect. Subjects reported instead a displacement along the path of observed motion. The occurrence of this so-called onset repulsion effect (Thornton 2002) is in agreement with the dynamic properties of the network model. Note that the spatial interactions within the network are omnidirectional. If the cooperative mechanisms are sufficiently strong there is a tendency for "extrapolation" also in the direction opposite to motion. The activity pattern that propagates beyond the point of onset does, however, not resonate with afferent inputs and thus quickly decays to resting level. Nonetheless, a readout of this pattern might explain the onset repulsion effect (see discussion in Thornton 2002).

In conclusion, we have shown that the proposed network model can place seemingly unrelated or even controversial findings about visual illusions observed with objects in motion in one coherent, unifying framework. With the recent advances in neuroscience we are beginning to understand the neural correlates of visual perception. Many systematic misperceptions remain, however, unsolved. They provide a fertile ground for combined modeling and physiological efforts that ultimately will lead to new insights into the complex structure of the visual system.

Acknowledgments

The work was supported by the BMBF (D. J.) and the Acções Integradas Luso-Alemas/DAAD. We would like to thank Beena Khurana and two anonymous reviewers for their helpful comments on a previous draft.

References

Abeles, M. (1991). *Corticonics*. Cambridge: Cambridge University Press.

Baldo, M. V. C., & Caticha, N. (2005). Computational neurobiology of the flash-lag effect. *Vision Res* **45**: 2620–2630.

Baldo, M. V. C., & Klein, S. A. (1995). Extrapolation or attention shift? *Nature* **378**: 565–566.

Ben-Yishai, R., Hansel, D., & Sompolinsky, H. (1997). Traveling waves and the processing of weakly tuned inputs in a cortical network module. *J Comp Neurosci* **4**: 57–77.

Berry II, M. J., Brivanlou, I. H., Jordan, T. A., & Meister, M. (1999). Anticipation of moving stimuli by the retina. *Nature* **398**: 334–338.

Bringuier, V., Chavane, F., Glaeser, L., & Frégnac, Y. (1999). Horizontal propagation of visual activity in the synaptic integration field of area 17 neurons. *Science* **283**: 695–699.

Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics* **27**: 183–228.

Dayan, P., & Abbott, L. F. (2001). *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems.* Cambridge, MA: MIT Press.

Douglas, R. J., Koch, C., Mahowald, M., Martin, K. A. C., & Suarez, H. (1995). Recurrent excitation in neocortical circuits. *Science* **269**: 981–985.

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science* **287**: 2036–2038.

Erlhagen, W. (2003). Internal models for visual perception. *Biol Cybern* **88**: 409–417.

Erlhagen, W., Bastian, A., Jancke, D., Riehle, A., & Schöner, G. (1999). The distribution of neuronal population activation (DPA) as a tool to study interaction and integration in cortical representations. *J Neurosci Methods* **94**: 53–66.

Erlhagen, W., & Jancke, D. (2004). The role of action plans and other cognitive factors in motion extrapolation: a modelling study. *Vis Cogn* **11**(2 /3): 315–340.

Fitzpatrick, D. (2000). Seeing beyond the receptive field in primary visual cortex. *Curr Opin Neurobiol* **10**: 438–443.

Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *J Exp Psych Learn Mem Cogn* **10**: 126–132.

Fröhlich, F. W. (1923). Über die Messung der Empfindungszeit. *Zeitschrift für Sinnesphysiologie* **54**: 58–78.

Fu, Y., Shen, Y., & Yang, D. (2001). Motion-induced perceptual extrapolation of blurred visual targets. *J Neurosci* **21** (RC172): 1–5.

Giese, M., & Xie, X. (2002). Exact solution of the nonlinear dynamics of recurrent neural mechanisms for direction selectivity. *Neurocomp* 44–46: 417–422

Gilbert, S. D. (1995). Dynamic properties of adult visual cortex. In M. Gazzaniga (ed.), *The Cognitive Neurosciences* (73–90). Cambridge: MIT Press.

Grinvald, A., Lieke, E., Frostig, R., & Hildesheim, R. (1994). Real-time optical imaging of naturally evoked electrical activity in intact frog brain. *J Neurosci* **14**: 2545–2568.

Grossberg, S. (1988). Nonlinear neural networks: principles, mechanisms, and architectures. *Neural Networks* **1**: 17–61.

Hahnloser, R., Douglas, R. J., Mahowald, M., & Hepp, K. (1999). Feedback interactions between neuronal pointers and maps for attentional processing. *Nat Neurosci* **2**(8): 746–752.

Hubbard, T. L. (1990). Cognitive representation of linear motion: possible direction and gravity effects in judged displacement. *Memory & Cognition* **18**(3): 299–309.

Hubbard, T. L., & Bharucha, J. J. (1988). Judged displacement in apparent vertical and horizontal motion. *Perception & Psychophysics* **44**: 211–221.

Hubbard, T. L., & Motes, M. A. (2002). Does representational momentum reflect a distortion of the length or the endpoint of a trajectory? *Cognition* **82**: B89–B99.

Jancke, D., Chavane, F., Naaman, S., & Grinvald, A. (2004a). Imaging correlates of visual illusion in early visual cortex. *Nature* **428**: 423–426.

Jancke, D., Erlhagen, W., Dinse, H. R., Akhavan, A. C., Giese, M., Steinhage, A., et al. (1999). Parametric population representation of retinal location: neuronal interaction dynamics in cat primary visual cortex. *J Neurosci* **19**: 9016–9028.

Jancke, D., Erlhagen, W., Schöner, G., & Dinse, H. R. (2004b). Shorter latencies for motion trajectories for flashes in population responses of cat visual cortex. *J Physiol* **556.3**: 971–982.

Khurana, B., & Nijhawan, R. (1995). Extrapolation or attentional shift? *Nature* **378**: 565–566.

Kirschfeld, K., & Kammer, T. (1999). The Fröhlich effect: a consequence of the interaction of visual focal attention and metacontrast. *Vis Res* **39**: 3702–3709.

Kreegipuu, K., & Allik, J. (2003). Perceived onset time and position of a moving stimulus. *Vision Res* **43**: 1625–1635.

Krekelberg, B., & Lappe, M. (2000). A model of the perceived relative positions of moving objects based upon a slow averaging process. *Vision Res* **40**: 201–215.

Krekelberg, B., & Lappe, M. (2001). Neuronal latencies and the position of moving objects. *TINS* **24**: 335–339.

Lamme, V. A. F., & Roelfsema, R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *TINS* **23**: 571–579.

Lennie, P. (1981). The physiological basis of variations of visual latency. *Vision Res.* **21**: 815–824.

Li, W., Piëch, V., & Gilbert, C. D. (2004). Perceptual learning and top-down influences in primary visual cortex. *Nat Neurosci* **7**(6): 651–657.

Maiche, A., Budelli, R., & Gómez-Sena, L. (2007). Spatial facilitation is involved in the flash-lag effect. *Vision Res* **47**: 1655–1661.

Metzger, W. (1932). Versuch einer gemeinsamen Theorie der Phänomene Fröhlichs und Hazelhoffs und Kritik ihrer Verfahren zur Messung der Empfindungszeit. *Psychologische Forschung* **16**: 176–200.

Müsseler, J., & Aschersleben, G. (1998). Localizing the first position of a moving stimulus: The Fröhlich effect and an attention-shifting explanation. *Perception & Psychophysics* **60**: 683–695.

Müsseler, J., Stork, S., & Kerzel, D. (2002). Comparing mislocalizations with moving stimuli: The Fröhlich effect, the flash-lag, and representational momentum. *Vis Cog* **9**(1/2): 120–138.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature* **370**: 256–257.

Nijhawan, R., Watanabe, K., Khurana, B., & Shimojo, S. (2004). Compensation of neural delays in visual-motor behaviour: no evidence for shorter afferent delays for visual motion. *Vis Cog* **11**(2/3): 275–298.

Purushothaman, G., Patel, S. S., Bedell, H. E., & Öğmen, H. (1998). Moving ahead through differential visual latency. *Nature* **396**: 424.

Salinas, E., & Abbott, L. F. (1994). Vector reconstruction from firing rates. *J Comput Neurosci* **1**: 89–107.

Thornton, I. M. (2002). The onset repulsion effect. *Spatial Vision* **15**: 219–243.

Whitney, D. (2002). The influence of visual motion on perceived position. *TICS* **6**: 211–216.

Whitney, D., Cavanagh, P., & Murakami, I. (2000). Temporal facilitation for moving stimuli is independent of changes in direction. *Vision Res* **40**: 3829–3839.

Whitney, D., Murakami, I., & Cavanagh, P. (2000). Illusory spatial offset of a flash relative to a moving stimulus is caused by differential latencies for moving and flashed stimuli. *Vision Res* **40**: 137–149.

Wilson, H. R., & Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik (Biol Cybern)* **13**: 55–80.

26

Perceiving-the-present and a unifying theory of illusions

MARK A. CHANGIZI, ANDREW HSIEH, ROMI NIJHAWAN,
RYOTA KANAI, AND SHINSUKE SHIMOJO

26.1 Introduction

Accurately perceiving where objects are in one's visual field is important for making decisions and interacting with the environment, but the visual system must contend with a significant delay – on the order of 100 msec (Lennie 1981; Maunsell & Gibson 1992; Schmolesky et al. 1998) – between the time of retinal stimulation and the time of the elicited percept. To deal with this delay, it has been hypothesized that the visual system has been selected to attempt to generate a percept that compensates for it, so as to perceive the present (Ramachandran & Anstis 1990; De Valois and De Valois 1991; Nijhawan 1994, 1997, 2001, 2002; Berry et al. 1999; Schlag et al. 2000; Sheth et al. 2000; Khurana et al. 2000; Changizi 2001, 2003, 2009; Changizi & Widders 2002). One circumstance where perceiving the present is crucial is when an observer is moving forward and approaching objects. It has been proposed that the classical geometrical illusion stimuli are due to fixations during forward motion and that the illusions are an expected consequence of perceiving the present mechanisms; that is, the classical geometrical stimuli are perceived not as they actually project but as they would project in the next moment if the observer were moving forward (Changizi 2001, 2003; Changizi & Widders 2002). This theory has been used to explain geometrical illusions such as the Hering, Orbison (Ehrenstein), Ponzo, Muller-Lyer, and Poggendorf. (See Appendix for a discussion of the distinction between projected size and distal size.)

Our main contribution here is the development and test of two kinds of prediction of this forward-motion perceiving-the-present hypothesis. The first prediction concerns the existence of dynamic versions of the classical geometrical illusions, where the converging lines are replaced by dynamic optic flow, and we show that such illusions exist as predicted (Section 26.3). The second prediction concerns a radical generalization of the theory, making a prediction about the pattern of illusions that should exist for twenty-eight different classes of stimuli (Section 26.4). In particular, we show that converging lines are just one of seven kinds of ecological cues to the observer's direction of motion (namely projected size, projected speed, luminance contrast, distance, eccentricity, converging lines, and optic flow itself) and that illusions are expected for three other modalities in addition to projected size or position (namely projected speed, luminance contrast, and distance). Via

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

a metareview of the literature we show that, consistent with the prediction, the theory unifies and systematizes more than 50 kinds of illusion into a 7 by 4 matrix of 28 classes of illusion. Before describing these two predictions we briefly review in Section 26.2 how the forward-motion perceiving-the-present hypothesis explains the classical geometrical illusions.

26.2 Review of how perceiving-the-present explains the classical geometrical illusions

Here we review how the forward-motion perceiving-the-present hypothesis has been used to explain the classical geometrical illusions.

26.2.1 The vanishing point of converging lines tends to indicate heading

The classical geometrical illusions are static, and, prima facie, one might not expect a dynamical hypothesis like perceiving-the-present to apply. However, consider that when an observer is in forward motion, the proximal stimulus at any time is typically due to a short "snapshot" fixation. Figure 26.1(a) illustrates an example instantaneous retinal stimulus during forward motion over a bridge. One can see that despite the picture being static, it is obvious in which direction the observer is heading. Importantly, notice that there are converging contours in the picture whose vanishing point is also the observer's direction of motion.

There are two sources of these converging contours. (1) They are due to real-world contours parallel to the observer's direction of motion, such as the sides of the road. In the carpentered worlds we tend to inhabit, there is a strong tendency for there to be such contours parallel to our movement (some may be seen in Fig. 26.1(a)), and thus a propensity for the vanishing point of converging contours to correlate with the observer's heading. (2) The converging contours in Fig. 26.1(a) are also due to optic flow itself, which engenders radial smear on the retina. The vanishing point of these "streak" lines is the focus of expansion (i.e., the point on the retina from which everything flows outward), which strongly correlates with the observer's heading either if the observer fixates on distant objects or if he fixates on approaching objects that tend to be near his direction of motion, something observers appear to do (Wann & Swapp 2000; Wilkie & Wann 2003).

Converging lines in a static picture like Fig. 26.1(a), therefore, suggest the observer's heading. Notice in this regard that it has long been noticed by cartoonists that streaks indicate trajectory, and that blurring an object suggests it is moving (in this regard see also Geisler 1999; Burr 2000; Ross et al. 2000; Geisler et al. 2001; Burr & Ross 2002; Cutting 2002). Streaks also activate MT (Krekelberg et al. 2003, 2005). (And, also, Kourtzi & Kanwisher 2000, have shown that static images depicting motion can activate MT.) As cues to observer motion, converging lines like those in the classical geometrical stimuli can be ambiguous (even though real motion streaks are not ambiguous), for backward motion would also engender converging lines. However, backward motion is ecologically rare, and the much

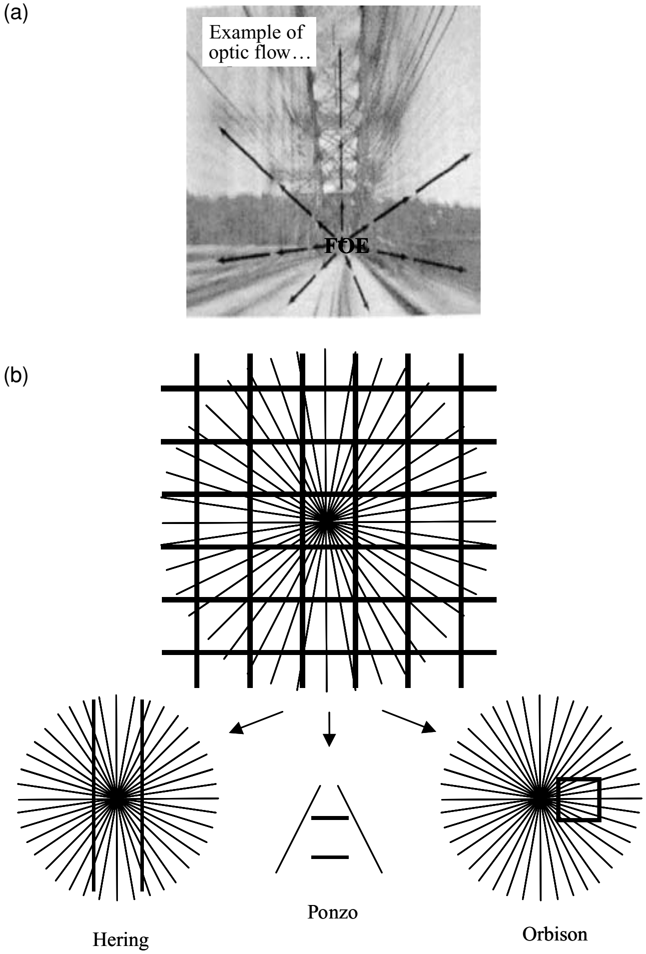


Fig. 26.1 (a) Illustration of the kind of stimulus received when in motion. It possesses many converging lines whose vanishing point coincides with the observer's direction of motion, and this is due to two reasons: (i) the fact that the contours along the road possess a vanishing point that the observer is heading toward, and (ii) that optic flow itself causes optic smear having a vanishing point that is the focus of expansion, which correlates strongly with observer heading. (b) Demonstration of the Hering, Orbison (or the Ehrenstein 1925), and Ponzo illusion. The Hering illusion is exemplified by the perceived curvature of the straight lines. The squares in the grid appear to be distorted, which is the Orbison illusion. And along the horizontal and vertical meridians, the line segments appear longer when closer to the center, which is a version of the Ponzo illusion. These three illusions are also shown by themselves below. The converging lines in these stimuli may provide a cue that the observer is heading toward the vanishing point, and many illusions follow from this assumption (Changizi 2001, 2003; Changizi & Widders 2002; Changizi et al., under review).

more probable direction of motion is forward (Tinbergen 1939, 1951; McBeath et al. 1992; Pavlova et al. 2002; Changizi & Widders 2002; Lewis & McBeath 2004). Converging lines can also, of course, be due to perspective. If forward-motion extrapolation mechanisms are enacted when the observer is not moving forward (e.g., if the converging lines are just due to perspective and there is no forward motion), this is not very costly compared to not perceiving-the-present when the observer is moving forward. Accordingly, it may be that there is little cost if compensation mechanisms are induced in the mere presence of converging lines, even when the observer is not moving.

26.2.2 Explaining the classical geometrical illusions

Now consider the classical geometrical illusions in Fig. 26.1(b) in light of the fact that converging lines may be due to an observer's forward movement. The forward-motion perceiving-the-present hypothesis states that, given the retinal stimulus, the visual system attempts to generate a perception that compensates for the approximately 100 msec delay between retinal stimuluation and elicited perception. For classical geometrical stimuli like those in Fig. 26.1(b), the hypothesis expects that the observer should perceive the stimulus not as it actually projects but how the scene would project in the next moment were the observer moving toward the vanishing point of the converging lines.

Consider the Hering illusion in Fig. 26.1(b) and let us ask how the two vertical lines would project to the eye in the next moment were the observer to move forward toward the vanishing point? In fact, the vertical lines on either side of the radial center would flow outward in the observer's visual field, and the parts of the vertical lines at eye level would flow outward most quickly. Visualize walking through a tall cathedral door. When far away, the sides of the door project roughly vertically (i.e., the angular distance between the sides of the door is roughly the same at all elevations). However, consider that when you get close to the cathedral door, the sides of the door above you converge toward one another toward the ceiling (i.e., the angular distance between the sides of the door at eye level is greater than far above your eye level). That is, the dynamics of projective geometry are such that when you move toward the center of two vertical contours, the contours actually do "bow out" in your visual field. In the Hering illusion this is exactly what one perceives. (This is not to be confused with perceiving the distal properties of the vertical lines to contort [see Appendix]). One always perceives the sides of the doorway to be straight real-world vertical lines while moving through it but perceives the projected positions of the sides of the doorway to bow out as described. It is these moment-by-moment changes of positions of objects in an observer's visual field that are in dire need of perceiving-the-present, for they change quickly over time, whereas distal properties such as the verticality of the sides of the doorway do not change over time.) Consider now a square to the right of the vanishing point like in the Orbison illusion (first noticed by Ehrenstein 1925). If the observer were moving toward the vanishing point, then in the next moment the left side of that square will have greater projected size than the right side of that square, which is how one perceives it in the illusion. This also explains a variant of the Ponzo illusion.

More generally, all the distortions one perceives in the larger grid are consistent with how the grid would project in the next moment were the observer moving toward the vanishing point. One can see, then, that if an observer were moving toward the vanishing point of the converging lines for the stimuli in Fig. 26.1(b), then the projections would vary in the next moment and be consistent with the way observers actually perceive these stimuli. (or "misperceive" because for these stimuli, the observer is not actually moving, and so the projections of the lines, grids, or squares do not actually change). This hypothesis has been applied more generally to explain many other classical geometrical illusions, including the Müller-Lyer, double-Judd, Poggendorff, corner, and upside-down "T" (Changizi 2001, 2003; Changizi & Widders 2002). (See these papers for conceptual criticisms of the traditional inference, or constancy-scaling explanation.)

26.3 Prediction of dynamic geometrical illusions

As discussed previously in our review of the explanation of the classical geometrical illusions, we supposed that the converging lines in the classical geometrical illusions may often be naturally encountered during a short fixation when an observer is moving forward, for optic flow will engender optic smear. Or, more weakly, our claim is that converging lines may trigger the same mechanisms as optic flow smear does. In this light, the visual system treats the classical geometrical stimuli as "snapshots" of intrinsically dynamic stimuli. If this hypothesis is true, it predicts the existence of *dynamic* classical geometrical illusions, where the converging lines are replaced by optic flow itself. In this section we carry out three tests of this prediction, using dynamic stimuli to modulate the observer's direction of forward motion (Experiment 2), to modulate the speed of forward motion (Experiment 3), and to modulate whether the observer is probably moving forward or backward (Experiment 4).

26.3.1 Modulating the observer's direction of forward motion

In the first dynamic experiment, we show that manipulating the direction of forward movement using a dynamic stimulus modulates perception as predicted. For two radially outflowing dots starting at the same elevation, Fig. 26.2(a) shows how their elevations differ in the next moment as a function of polar angle around the observer's direction of motion. For example, if both dots are in the observer's upper right quadrant as shown in the illustration on the left of Fig. 26.2(a), then in the next moment the left dot will rise higher than the right dot, and this is shown in the upper right quadrant of the "horizontal dots" plot (positive illusion magnitude meaning that the inner dot undergoes greater vertical angular displacement in the next moment). The "vertical dots" plot is analogous, but where dots are above one another, it concerns horizontal displacement. The two plots in Fig. 26.2(a) amount to predictions.

The "horizontal dots" plot in Fig. 26.2(b) shows how observers perceive the relative elevations of two side-by-side dots in a radial display, as a function of polar angle around

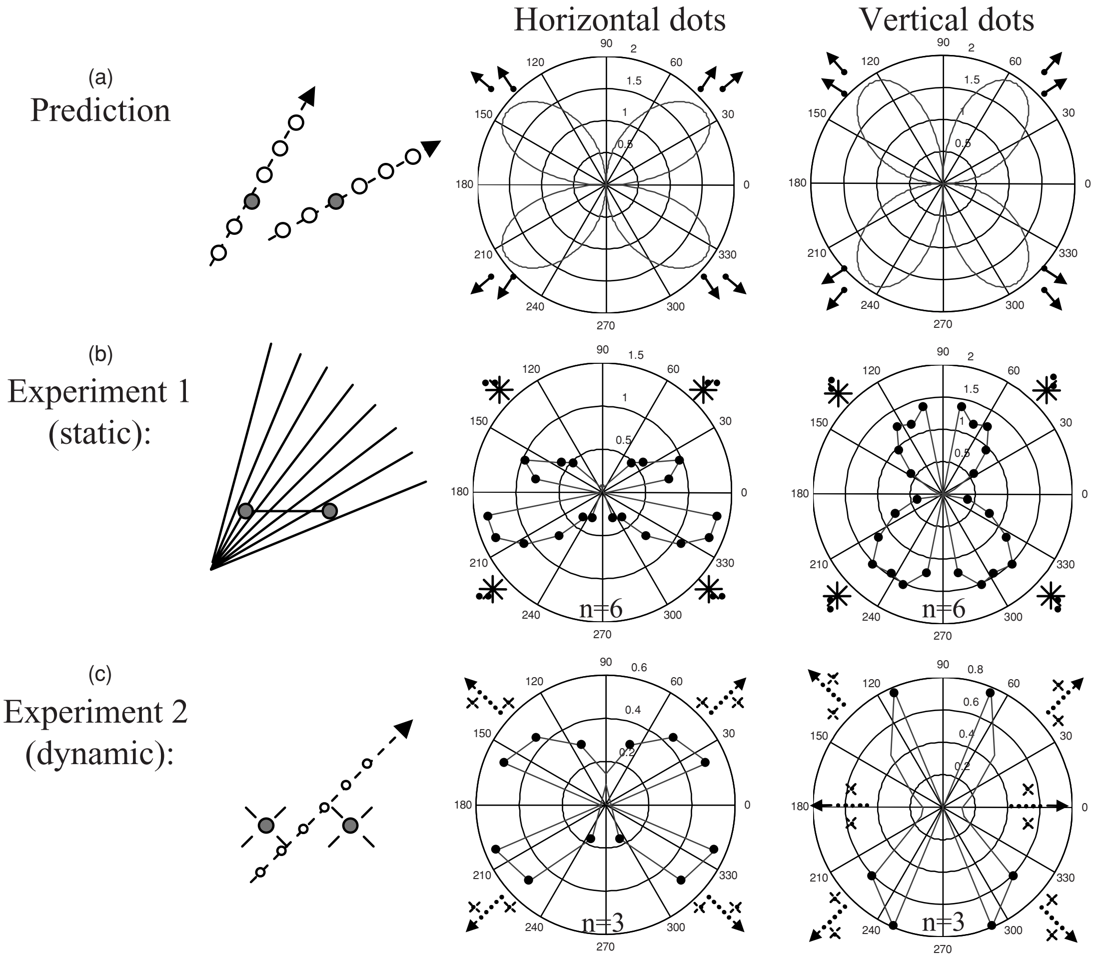
Horizontal dots
Vertical dots
(a)
Prediction
(b)
Experiment 1
(static):
(c)
Experiment 2
(dynamic):
90
60
30
0
330
300
270
240
210
180
150
120
2
1.5
1
0.5
0.6
0.4
0.2
0.8
n=6
n=3

Fig. 26.2 Predicted illusion magnitude (a) and measured illusions for static (b) and dynamic (c) illusions. (a) The predicted misperception as a function of polar angle around the direction of motion. On the left is a figure conveying the fact that when the two optically flowing dots have the same projected distance from the horizontal meridian (the gray dots), the one nearer the direction of motion will, in the next moment, typically project farther from that meridian than the other dot. The "horizontal dots" polar plot in the middle of this row shows how much farther the near-direction-of-motion dot moves away from the horizontal meridian than the far-from-direction-of-motion dot in the next moment, measured as the projected angle made between the horizontal meridian and the imaginary projected line connecting the dots. These predicted values are computed as follows: The midpoint between the pair of horizontal dots is placed on a 1 m radius circle in the forward-moving observer's frontoparallel plane, at a distance of 1 m in front of the observer. The dots themselves are placed 0.5 m on either side of the midpoint. The observer is assumed to move at 1 m/sec, and the amount of vertical displacement is computed over a (latency) time interval of 100 msec. These values were chosen somewhat arbitrarily, and the qualitative predictions are not dependent on these values. Analogous computations were made for the "vertical dots" plot, which is similar, but where the dots are above one another and the issue concerns the next-moment projected distance of the dots from the vertical meridican. (b) and (c): results for Experiments 1 and 2. The two illustrations on the left in the following two rows describe stimuli that are plausibly due to optic flow projection dynamics of the kind just described in row (a), and each has two gray dots at identical projected distance from the horizontal meridian. (b) is for the static, classical geometrical illusions, and (c) is for a dynamic version of the classical geometrical illusions. Because the near-direction-of-motion dot is, in the next moment, going to be shifted farther away from the horizontal meridian, perceiving-the-present accordingly expects observers to perceive the near-direction-of-motion dot to be shifted in this way. The method of adjustment was used for each experiment. For each of these kinds of stimulus, the averaged results are shown for where the dots are horizontally aligned and where the dots are vertically aligned. In each case the expected illusions exist in each quadrant, showing substantial similarity to the prediction in (a). Illusions are measured in degrees of slant, where the illusion is positive if the dot nearer to the observer's direction of motion is perceived farther from the meridian than the other dot. Number of subjects, n, is shown in each plot. Corresponding positions in the left and right half of the visual fields have been averaged together, so the left and right side of each plot are identical and redundant. Dots at a point on the graph indicate that the point is significantly greater than zero at the $p < 0.05$ level (via t-test), where an observer's responses on the left side were treated as independent of his responses on the other side (the degrees of freedom is therefore twice the number of subjects minus two).

the radial display (Experiment 1). For example, for the stimulus shown on the left of Fig. 26.2(b), the empirical illusion magnitude is shown in the upper right quadrant of the "horizontal dots" plot, which means that the left dot is perceived to be higher than the right dot. Note that this is consistent with how the elevations of the two dots will change in the next moment, as predicted by Fig. 26.2(a). That is, the two plots in Fig. 26.2(b) for static radial line illusions fits the prediction of the forward-motion perceiving-the-present hypothesis.

In Experiment 2 the static converging lines of Experiment 1 are replaced by a flowing dot (6 deg/sec), and the two target dots (separated by 5 deg) are briefly flashed (for about 40 msec) just as the flowing dot passes them, illustrated by the figure on the left of Fig. 26.2(c). The plots in Fig. 26.2(c) show the average illusion magnitude across all the conditions, and one can see that this dynamic classical geometrical illusion is modulated by direction of motion as predicted (Fig. 26.2(a)), and has the same signature as the static geometrical illusion in Fig. 26.2(b).

26.3.2 Modulating the observer's speed of forward motion

We have just seen that modulating the inferred observer direction of motion via dynamic stimuli modulates perception as expected and analogous to static geometrical illusions. Here we demonstrate that increasing the inferred speed of forward movement via dynamic optic flow increases misperceptions as expected. In the static domain, this is analogous to adding more converging lines (because when moving faster, a greater number of objects tend to be moving sufficiently fast to induce optic blur), and it is well known that classical geometrical illusions such as the Ponzo are stronger when there are more converging lines.

Figure 26.3 illustrates the basic stimulus design, where the bottom hemifield possesses optically flowing dots, and two Ponzo bars are briefly flashed. There were two versions of the flowing dots, "slow" and "fast." Here we expect, and find, that the upper Ponzo line should be perceived as projecting larger, because it is nearer to the observer's direction of motion. Points of subjective equalities were computed for each observer on the "slow" and "fast" conditions: the average illusion was 3.9% in the "slow" condition, and 5.6% in the "fast" condition, where a positive value indicates the upper Ponzo line was perceived to project larger than the lower Ponzo line. The amount of illusion was significantly greater in the "fast" condition: the average difference was 1.7% (standard error = 0.52), and this is significantly above 0 by a t-test ($p = 0.0101$, $t = 3.24$, $df = 9$). On an observer-by-observer basis, nine of the ten observers perceived a greater illusion in the "fast" condition than the "slow" condition.

26.3.3 Forward versus backward

We have thus far seen in this section that, as predicted, dynamic classical geometrical illusions exist, and manipulation of observer forward direction (Experiment 2) and speed

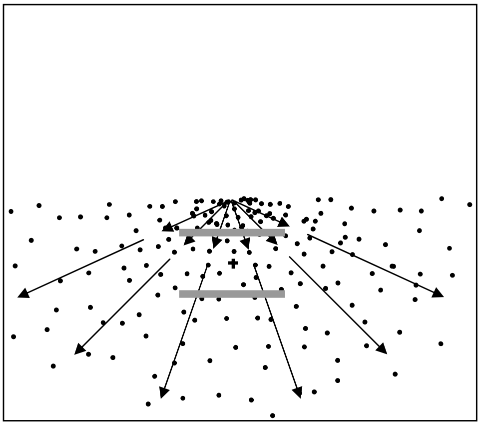


Fig. 26.3 Illustration of an experiment (Dynamic Ponzo-Flash Illusion, Experiment 3) demonstrating that the speed of optic flow modulates illusions as expected. The arrows here indicate that the black dots flowed outward with a velocity gradient. The observer fixated on a red dot (shown here as a black cross) between the two Ponzo lines, and two red Ponzo lines (shown here as gray) were briefly flashed. Faster optic flow, consistent with faster forward movement, led to greater illusions (i.e., greater misperceptions that the upper bar has larger angular size than the lower bar). Optically flowing black dots were presented in the lower half of a white screen, simulating forward motion toward the focus of expansion of those dots. Dots flowed only in the hemifield below the focus of expansion. Two simulated observer speeds were used, "slow" and "fast": dots began with angular speed 0.80 deg/sec in "slow" and 1.60 deg/sec in "fast" condition; and acceleration 5.39 deg/sec2 in "slow" and 10.78 deg/sec2 in "fast." These values were chosen because they led to qualitatively different simulated observer speeds. The observer fixated on a red dot (shown in Fig. 26.3 as a black cross) 3.34 deg below the focus of expansion. After 3 sec, two red, horizontal, Ponzo line segments briefly flashed (0.036 sec) above and below the fixation point (1.90 deg above and below), each below the focus of expansion. In the two-alternative forced-choice design, the projected lengths of the upper and lower Ponzo line segments varied over 7 pairs of values, ranging from upper segment being 6% longer than lower segment to lower segment 12% longer than upper segment (with center around 4.6 deg or arc). In total, then, there were seven kinds of Ponzo-line presentations, and two optic-flow speed conditions, for a total of fourteen distinct stimuli. Each of these was presented ten times, randomly interleaved. After each presentation of the flow followed by the flashed Ponzo lines, observers judged whether the upper or lower Ponzo line appeared larger in projected length. The experiment took about 20 min to complete. Ten observers (two nonnaïve, eight naive) participated in the experiment. All observers had normal or corrected-to-normal visual acuity.

(Experiment 3) modulates the illusions as expected. It is useful to mention one kind of experiment that may be deemed an appropriate test of our theory: backward motion (see Changizi & Widders 2002, for discussion of this). The reader may wonder, for example, if optical contraction consistent with backward motion should lead to a counter-Ponzo illusion in the dynamic Ponzo-flash experiment. However, we feel that this type of experiment may only be a weak test of our theory. The visual system can be expected to be competent at perceiving-the-present only under conditions sufficiently similar to the natural conditions of stimulation, either evolutionarily or during the animal's lifetime. Backward motion is quite infrequent. Furthermore, when moving backward one is not at risk of colliding into objects in one's view. These facts together suggest that there is less selective pressure for perceiving-the-present mechanisms for backward motion. There is therefore reason to doubt that the visual system might be able to correct for latencies under backward-motion conditions.

Nevertheless, it would be interesting to see whether the visual system can sometimes appropriately respond to backward-motion stimuli. Experiment 4 describes a simple kind of backward-motion stimulus, but one that nevertheless possesses abundant cues indicating that the target objects are receding away from the observer. Abundant cues are probably necessary for the "backward" interpretation because of its relative infrequency (Lewis & McBeath 2004), similar to the bias against a "backward" interpretation of other objects (Tinbergen 1939, 1951; McBeath et al. 1992; Pavlova et al. 2002).

This experiment consisted of three conditions: "stationary control," "backward," and "forward." See Fig. 26.4(a) for illustration. In each condition, each trial presented two frames in short succession, and in each frame (except the control) there were four vertical line segments, horizontally aligned – two to the left of fixation (left pair) and two to the right (right pair). In each condition, the lines in the second frame are identical (and are akin to the Ponzo, but one on each side of the center, and without the converging lines); only the first frames differ, and differ so as to indicate backward motion, forward motion, or to serve as a control. Observers were asked to judge whether the inner lines or the outer lines in the second frame are larger. In the stationary control condition, observers perceived the inner pair of lines to be 4.3% larger (95% confidence interval (3.9, 4.7), via standard bootstrap) than the outer pair, consistent with the effects of eccentricity alone (see case 5A of Table 26.2). In the forward condition, the illusion was significantly amplified to 8.2% (95% confidence interval (7.6, 8.8)). In the backward condition, the illusion was significantly diminished compared to the control, namely 1.2% (95% confidence interval, [0.8, 1.7]). That is, compared to the control, forward and backward motion modulate the illusion in the expected directions (Fig. 26.4(b)).

26.3.4 Summary

In Section 26.2, we reviewed how a forward-motion perceiving-the-present hypothesis can explain many classical geometrical illusions. The hypothesis predicts that the same illusions should be elicited if, informally, the converging lines are replaced by dynamic optic flow. In this section, we showed that this is the case. We showed illusions like those in the classical geometrical ones can be made to occur, but using dynamic optic flow stimuli, where we modulated the inferred observer direction of motion (Experiment 2) and speed (Experiment 3); we even showed conditions under which backward motion can induce the expected counterillusion (Experiment 4).

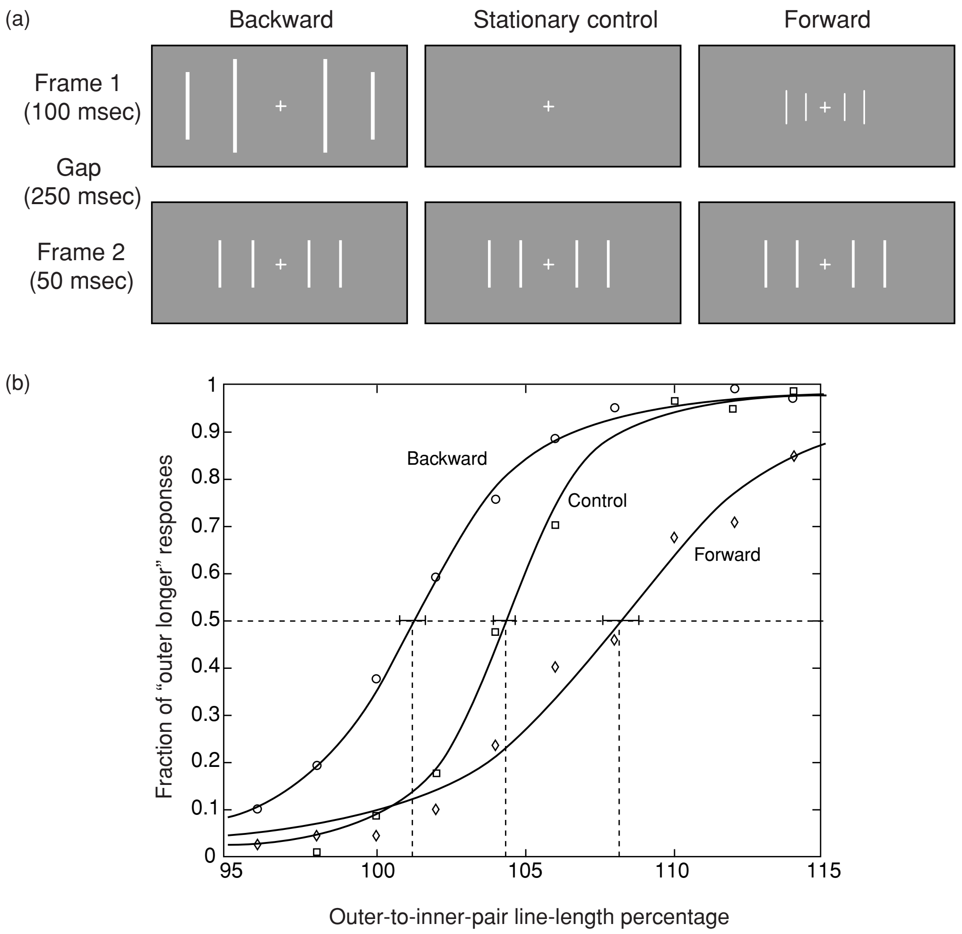


Fig. 26.4 (a) Illustration of the three conditions of Experiment 4, shown in the three columns here. The two rows show the first and second (which is the last) frame of the stimuli. The second frames are identical in all three conditions, and the outer lines of the second frame are varied in the two-alternative forced choice design. In the "stationary control" condition, the first frame has no lines at all. In the "backward" condition, the lines in the first frame are arranged approximately consistent with how they would project if, in the next frame, they receded away from the observer and projected as in frame 2. And in the "forward" part, the lines in the first are arranged approximately consistent with how they would project if, in the next frame, they approached the observer and projected as in frame 2. Observers were required to judge whether the inner lines or the outer lines in the second frame are larger. Ten trials were performed for each setting of the outer line segment's length, which varies over eleven values. Twelve (three nonnaïve, nine naive) observers participated in this experiment. (b) Results of Experiment 4, for all 12 observers. The plot shows, for the three conditions, the fraction of "outer is longer" responses versus the relative physical settings of the outer and inner lines (measured here as the outer-to-inner percentage, 100% meaning the outer and inner are the same projected size). In the stationary control condition (squares), observers had to increase the size of the outer pair to 4.3% larger (95% confidence interval (3.9, 4.7), via standard bootstrap) than the inner pair to perceive them as equal. That is, observers perceived the inner lines to be 4.3% larger than the outer lines. In the forward condition (diamonds), the illusion was significantly enhanced to 8.2% (95% confidence interval (7.6, 8.8), whereas in the backward condition (circles) the illusion was significantly diminished to 1.2% (95% confidence interval (0.8, 1.7).

26.4 Predicted table of 7 × 4 = 28 illusion classes

In Section 26.3, we presented evidence confirming the prediction of our forward-motion perceiving-the-present hypothesis that there should be dynamic versions of the classical geometrical illusions, where dynamic flow replaces the static converging lines. In this section we present a second prediction of the hypothesis, one concerning a predicted pattern over a large swath of illusions from the visual perception literature.

The idea behind this new prediction is a generalization of the forward-motion perceiving-the-present idea used to accommodate the classical geometrical illusions and discussed in Section 26.2. The explanation of the classical geometrical illusions given there required (i) using vanishing point cues in the stimulus to determine the probable observer's direction of motion, and (ii) working out how the projected sizes of objects in the scene will change in the next moment when the observer moves in that direction, which depends on where the objects are in the visual field relative to the direction of motion (e.g., the Hering lines bow outward in the visual field more quickly at eye level). More generally, we wish to look for (I) cues to the observer's direction of motion, and (II) the rates at which properties tend to change depending on where they are in the visual field relative to the observer's direction of motion. In the following two subsections we discuss two kinds of optic flow regularities, concerning (I) and (II) respectively.

26.4.1 Optic flow regularity type I: correlates of direction of motion

We first describe the correlates of the observer's direction of motion. Fig. 26.5(a) and 26.5(b) are photographs taken while in forward motion, and the observer's direction of motion is obvious, for there are many cues for it. Many of the correlates of the direction of motion can be understood by examination of Fig. 26.5, and we enumerate them below.

A region of the visual field nearer to the observer's direction of motion tends to have

(1) Smaller projected sizes.
(2) Smaller projected speeds.
(3) Greater luminance contrasts.
(4) Greater distances from the observer.
(5) Lower eccentricity.
(6) Lower projected distance from the vanishing point of converging lines.

These six correlates are recorded in Fig. 26.5(c). Notice that correlate (6) is just the correlate mentioned earlier in Section 26.2 concerning the classical geometrical illusions.

Although Fig. 26.5 provides examples of forward-moving scenes, the most fundamental reason for these correlates is this: *When one moves forward, one must avoid obstacles lest one collide with them.* When moving forward, the direction of motion is therefore different than other places in the visual field, for the direction of motion must have some "room for forward movement." That is, the distance must be sufficiently great for some degree of forward movement. The other places in the visual field, however, are not under any such constraint: they can be near or far. That is, regions of the visual field nearer to the direction

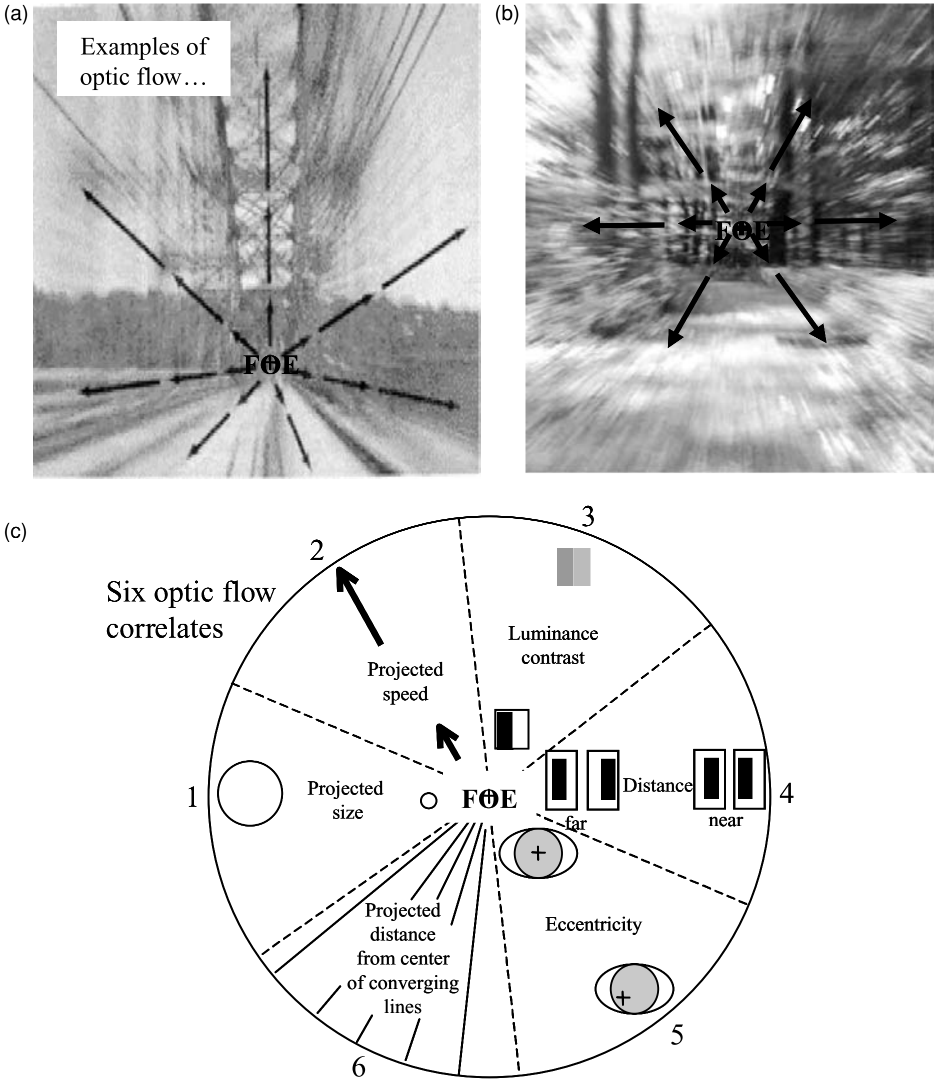


Fig. 26.5 (a) This picture (from the public domain and also shown in Fig. 26.1(a)) illustrates many of the correlates of optic flow and the direction of motion. Namely, moving from the direction of motion (the focus of expansion) outward, projected sizes increase (e.g., the road), projected speeds increase (the arrows), luminance contrasts decrease (notice the overhead structures), distance decreases, and projected distance from the vanishing point of converging lines increases. (b) Another picture (from the public domain), this one of optic flow in a forest. (c) The circle signifies an observer's visual field, and the center the location of the focus of expansion (FOE), and the observer direction of motion. Around the circle are shown 6 correlates of optic flow, labeled 1 through 6. For example, correlate 1 is for projected size, and tells us that projected sizes are smaller near the observer direction of motion and get larger farther from the observer direction of motion. Some of the correlate descriptions need comment. Distance can be cued via many sources of information, but the distance correlate is shown here via using two stereograms, intended for uncrossed viewing: the one on the left depicts a single black bar behind a rectangular frame (i.e., distance of the black bar is great), and the stereogram on the right depicts a single black bar in front of a frame (i.e., the black bar is near). All but correlate 5, eccentricity, are exemplified by (a) and (b). The eccentricity correlate is due to the fact that observers are typically looking in the direction they are headed, and in (c), this is signified by an eye with a cross at a location on the retina.

of motion will correlate with being farther from the observer. This argument is very general and would even apply, say, for a rocket ship moving within an asteroid field, where there is no ground plane. In the real world that we inhabit, there is a ground plane, and very often walls and ceiling, and in these circumstances the correlation between direction of motion and distance from the observer is even stronger. For the reasons just mentioned, correlate (4) follows. What about the other five correlates? We already discussed correlate (6) in Section 26.2 when we reviewed the perceiving-the-present explanation of the classical geometrical illusions. Correlate (1) follows from correlate (4) because as an object nears the observer, its projected size increases (i.e., nearer objects have greater projected size). Because distance and projected size are independent of an observer's pattern of fixation, correlates (1) and (4) are true no matter the manner in which a moving observer retinally tracks. The other correlates depend on the observer's pattern of fixation, however. We now consider the two possible cases.

Case 1, a constant angle of fixation relative to the direction of motion: If an observer's fixation is at some constant angle relative to the direction of motion – that is, the observer does not fixate on approaching objects – then correlates (2), (3), and (6) result for the following reasons. Correlate (2) follows from correlate (4) because as an object nears the observer, it will also tend to be a greater projected distance from the direction of motion, and its projected speed will increase. Correlate (3) follows, in turn, because luminance contrast and projected speed are inversely related. To see why, consider an object flowing across a one-degree-long segment of the projection sphere. The luminance contrast at that one-degree-long segment is just (roughly) the magnitude of the difference in luminance between it and that of the background luminance. Objects with greater projected speed integrate along the one-degree-long arc for a shorter period of time, and thus the luminance of that arc – being so "smeared" – will differ less from that of the surround (or background) luminance. See Section 26.2 for a discussion of correlate (6). Correlate (5), though, may not hold if one is fixated at a large angle from the direction of motion. However, if one makes the reasonable assumption that forward-moving observers have a tendency to look roughly in the direction they are going – something argued to be optimal (Wann & Swapp 2000) and for which there is evidence that people do (Wilkie & Wann 2003) – correlate (5) does follow. So, if observers tend to fixate at some constant angle relative to the direction of motion, and if they tend to look roughly where they are going, then all the correlates follow.

Case 2, fixating on approaching objects: However, observers often fixate on approaching objects rather than fixating at some constant angle relative to the direction of motion. When fixated on an approaching object, the focus of expansion of outflowing dots will be at the point of fixation, not the direction of motion (Regan & Beverly 1982). The fixated point, and not the direction of motion, will then tend to have smaller projected speeds, greater luminance contrasts, lower eccentricity, and lower projected distance from the vanishing point of converging lines. That is, correlates (2), (3), (5), and (6) will not necessarily hold. However, as mentioned above, forward-moving observers tend to look where they are going – that is, lower eccentricity tends to correlate with heading – and the focus of expansion consequently tends to covary with the direction of motion, and (2), (3), (5), and (6) therefore *do* follow (and (1) and (4) were independent of fixation patterns).

In summary, we have just derived that correlates (1) through (6) are true for forward-moving observers no matter their fixation patterns (i.e., no matter whether Case 1 or 2), so long as they tend to look approximately where they are going.

Several observations are important to mention. (i) One must recognize that even though "projected sizes tend to be smaller near the direction of motion," it does not follow that every stimulus with a projected size gradient is a stimulus that would be naturally encountered while the observer is in motion. That is, these ecological regularities tell us that ecologically natural optic flow stimuli have certain characteristics (like a projected size gradient), but they do not tell us that any stimulus with these characteristics is ecologically natural. A similar point of caution holds for correlates (2), (3), and (4) as well. For example, although ecologically natural optic flow stimuli have lower projected speeds near the direction of motion, consider a stimulus with lower projected speed objects in one part of the stimulus but where the objects have random directions. Such a stimulus may not be ecologically associated with optic flow. (ii) Note that these ecological regularities do *not* require an assumption of carpentered environments (and recall that converging lines may typically be due to optic smear). (iii) Note that correlate (3) implies that, when an observer is in motion, nearer objects tend to be lower in luminance contrast, which is in contradistinction to the weaker ecological regularity governing when an observer is not moving, where nearer objects tend to have greater luminance contrast.

26.4.2 Optic flow regularity type II: how quickly features change depending on nearness to the direction of motion

With the six ecological correlates of direction of motion now enumerated, we must discuss another kind of ecological regularity, one concerning the *rates* at which change occurs as a function of projected distance from the direction of motion. (By "projected distance from the direction of motion" we mean the visual angle between the observer direction of motion and some object in the visual field.) When an object is near to passing you and the projected distance from the direction of motion is accordingly great, its projected size and speed have nearly asymptoted to their maxima, its luminance contrast has nearly reached its minimum (because it varies inversely with speed), and its distance from the observer has reached its minimum. Said differently, projected size, projected speed, luminance contrast, and distance (from the observer) undergo little change when close to ninety degrees from the observer's direction of motion; these features undergo their significant changes when nearer to the direction of motion.

It is possible to derive the following ecological regularities concerning the rates of growth:

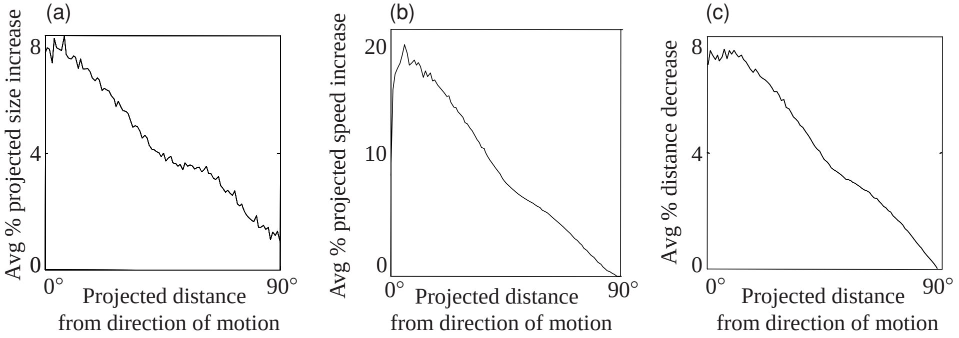


Fig. 26.6 (a) The average percentage projected size growth as a function of projected distance from the direction of motion. It was obtained by simulating 10^6 forward movements at 1 m/sec for 100 msec and computing the average percentage projected size growth of a line segment of random length (between a few centimeters and a meter), orientation, and placement (no more than 2 meters to the side, above, below, or in front of the observer's eye). We have confined objects to be relatively near the observer because we believe that it is the dynamics of nearby objects that will have tended to shape the functions computed by the visual system. The qualitative shape of the plots does not change if the boundaries of the simulated world are scaled up uniformly. (b) and (c) are similar to (a), but recording, respectively, the average percentage projected speed increase and the average percentage distance decrease, each as a function of projected distance from the direction of motion. In sum, for relatively nearby objects, projected size, projected speed, luminance contrast, and distance undergo greater percentage change when nearer to the direction of motion.

For two objects of similar distance from passing the observer, the one nearer to the observer's direction of motion undergoes, in the next moment,

(A) a greater percentage increase in projected size,
(B) a greater percentage increase in projected speed,
(C) a greater percentage decrease in luminance contrast, and
(D) a greater percentage decrease in distance from the observer.

These qualitative intuitions can be made rigorous and quantitative by simulating forward movement. Projected sizes tend to increase most when near the direction of motion, to increase by about half that when at 45 deg from the direction of motion, and to not increase at all when passing the observer (Fig. 26.6(a)). Similar conclusions hold for projected speed (and thus luminance contrast) and distance (Fig. 26.6(b) and 26.6(c)). In our simulations we assume that the objects are "relatively nearby," and specifically no more than 2 m to the side, above, below, or in front of the observer's eye. This "relatively nearby" assumption is reasonable for two reasons. First, the objects where latency compensation is most needed are the ones near enough to interact with; there will be little or no selection pressure for the compensation of objects, say, 100 m distant from an observer (Cutting & Vishton 1995). Second, most objects very far away will simply be too small to notice, and, furthermore, any changes they undergo will be small in absolute magnitude and thus insignificant compared to the changes of nearby objects. (Note that this "relatively nearby" was not made for optic flow regularity I concerning the six correlates of the observer's direction of motion, because even far away, unchanging stimulus features – despite not requiring compensation – can provide information concerning the observer's direction of motion.) However, these conclusions – and correlates (A) through (D) – are general; for example, increasing the 2 m limit to some larger value does not modify the shape of the plots. The main qualitative result is due to the simple fact that, as mentioned above, objects near to passing you are no longer undergoing percentage change in projected size, projected speed, luminance contrast, and distance.

26.4.3 Twenty-eight distinct ecological regularities

We have now introduced two broad kinds of ecological optic flow regularity: (I) correlates-of-direction-of-motion, and (II) how-quickly-features-change-nearer-the-direction-of-motion.

Within the first kind we introduced six correlates of the direction of motion ((1) through (6) from above), and within the second kind we introduced four features that change more quickly in the next moment when nearer to the direction of motion ((A) through (D) from above). These two kinds of optic flow regularity are robust, qualitative, statistical generalizations and do not rely upon any posthoc setting of parameters. It is important to understand that, although these two kinds of regularity are related, they are very different and are independent of each other. Given a stimulus, the first group of regularities ((1) through (6)) helps us to determine, from the stimulus, the observer's direction of motion. These six play the same role as the converging lines did in the classical geometrical illusions. We argued in Section 26.2 that the vanishing point of converging lines is probably the observer's direction of motion. This converging-line regularity is now just one of six such regularities. Once we have used the first group of regularities to determine the observer's direction of motion – for example, the observer is moving toward the vanishing point of the converging lines – we then need to determine how features will change in the next moment were the observer to move in that direction.

The second group of regularities ((A) through (D)) tells us how features change in the next moment; the rate at which features change depends on where they are in the visual field relative to the observer's direction of motion. On average, nearby objects closer in the visual field to the observer's direction of motion will undergo greater change in the next moment (i.e., the derivative is steep nearer the direction of motion). And this is what the second group of regularities told us. For example, for the Orbison illusion as shown in Fig. 26.1, given that the observer is moving toward the vanishing point (something determined via the first group of regularities), we want to know how the projected nature of the square will change in the next moment. This latter issue is answered via knowing that the projected sizes of objects tend to increase more in the next moment when they are nearer to the observer's direction of motion, and so the left side of the square in the Orbison illusion will grow more in the next moment than the right side. (And importantly, the left and right side of the square are probably not too different in distance from passing the observer; namely, in this case they probably lie in the observer's frontoparallel plane and so are equally distant from passing the observer.)

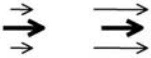
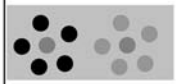
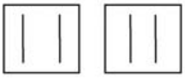
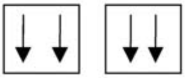
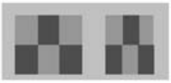
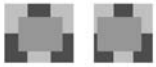
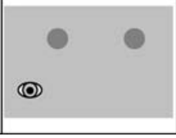
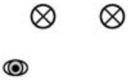
Together, these two kinds of ecological regularity tell us which parts of the visual field will undergo greater feature changes in the next moment. In particular, from the six correlates-of-direction-of-motion regularities and the four how-quickly-features-change-nearer-to-the-direction-of-motion regularities, one can distinguish between $6 \times 4 = 24$ distinct ecological regularities. Consider, for example, combining (1) and (B) (we shall call such a combination "1B"). This combination determines a specific ecological regularity, namely (1) that a region of the visual field with lower projected sizes tends to be nearer the direction of motion, and (B) for two objects of similar distance from passing the observer, the one nearer the direction of motion tends to undergo, in the next moment, a greater percentage increase in projected speed. From this we may infer the following more succinct statement of 1B: *For two objects of similar distance from passing the observer, the one nearer the region of the visual field with smaller projected sizes tends to undergo, in the next moment, a greater percentage increase in projected speed.* Consider, as another example, combining (3) and (C) to make ecological regularity 3C: (3) a region of the visual field with greater luminance contrasts tends to be nearer the direction of motion, and (C) for two objects of similar distance from passing the observer, the one nearer the direction of motion tends to undergo, in the next moment, a greater percentage decrease in luminance contrast. Again, it is reasonable to expect that the following shorter statement of 3C is true: *For two objects of similar distance from passing the observer, the one nearer the region of the visual field with greater luminance contrasts tends to undergo, in the next moment, a greater percentage decrease in luminance contrast.*

Table 26.1 is a matrix showing all twenty-four of these ecological regularities, with the correlates of the direction of motion as the rows and, as the columns, the four features that change more quickly when nearer to the direction of motion. The table also includes a seventh row, where the correlate of the direction of motion is the focus of expansion of optic flow itself (such stimuli tend to possess more than one of the six stated correlates of the direction of motion). In total, then, Table 26.1 catalogs twenty-eight distinct ecological regularities relating disparate stimulus types to four modalities of perception.

26.4.4 Twenty-eight distinct predicted illusion classes

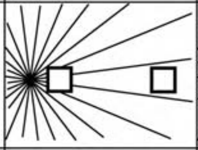
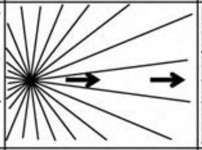
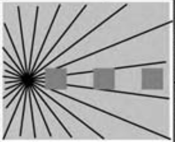
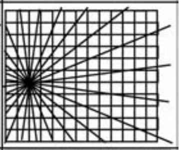
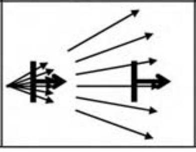
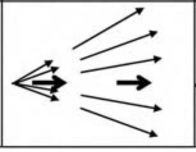
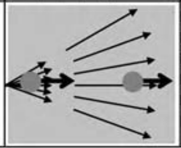
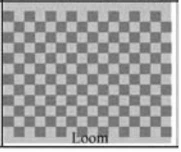
What do these ecological regularities have to do with visual perception? These twenty-eight distinct ecological regularities are important because they *also* amount to twenty-eight distinct predicted illusion classes. This is because, under perceiving-the-present, the perception is predicted to be representative of the way the scene will be in the next moment (i.e., by the time the perception occurs). Each ecological regularity in Table 26.1 states how features will change in the next moment, and perceiving-the-present therefore expects observers to have perceptions that accord with these expected next-moment features. More specifically, the predicted illusions recorded in Table 26.1 can be described as follows: For each class there are two similarly distant target objects that are identical in regard to the

Table 26.1 *Prediction of the optic flow regularities hypothesis. Catalog of the twenty-eight ecological correlates of forward motion and the twenty-eight illusion classes from the perceiving-the-present framework due to the effects of seven direction-of-motion correlates (the rows) on perceived projected size, projected speed, luminance contrast, and distance (the columns). To illustrate how to read the table, the following is how the upper left case of the table, illusion class 1A, should be read: "A region of the visual field with lower projected sizes (greater projected spatial frequency) is associated with, in the next moment (i.e., the predicted perception is of), a greater increase in projected size (greater decrease in projected spatial frequency)." Each square also shows an example figure consisting of (a) two targets that are identical in the modality of the column, but (b) differ with respect to the feature of the row. The probable direction of observer motion is always toward a point on the left side. For Row (2) and Column (B) – each of which concern motion – arrows are used to indicate stimulus speed and direction. For Row (4) stereograms (meant for divergent viewing) are used for the example figures; although we have used stereo disparity to cue relative distance, any cue to relative distance could be used. For Row (5), the little eye in the figures represents the approximate fixation point.*

For two objects of similar distance from passing the observer, the one nearer the region of the visual field with… ↓	**…tends to undergo, in the next moment (i.e., the predicted perception is of)…**			
	(A) …a greater increase in (angular) size (i.e., larger on left)	(B) …a greater increase in (angular) speed (i.e., faster on left)	(C) …a greater decrease in luminance contrast (i.e., lower contrast on left)	(D) …a greater decrease in distance (i.e., nearer on left)
(1) …lower (angular) sizes…				
(2) …lower (angular) speeds…				
(3) …greater luminance contrasts…				
(4) …greater distances…				
(5) …lower eccentricity…				

(*cont.*)

Table 26.1 (*cont.*)

For two objects of similar distance from passing the observer, the one nearer the region of the visual field with… ↓	**…tends to undergo, in the next moment (i.e., the predicted perception is of)…**			
	(A) …a greater increase in (angular) size (i.e., larger on left)	(B) …a greater increase in (angular) speed (i.e., faster on left)	(C) …a greater decrease in luminance contrast (i.e., lower contrast on left)	(D) …a greater decrease in distance (i.e., nearer on left)
(6) …lower (projected) distance from vanishing point…				
(7) …lower (projected) distance from focus of expansion…				

column modality. The region of the visual field near the left target is given the features specified by the row, and this thereby makes it probable that the left region is nearer to the observer's direction of motion. The target object on the left is therefore predicted to be perceived by observers to have a column modality that changes in the way stated in the column heading.

For example, ecological regularity 1A states that *for two objects of similar distance from passing the observer, the one nearer the region of the visual field with smaller projected sizes tends to undergo, in the next moment, a greater percentage increase in projected size*. Perceiving-the-present accordingly predicts that, when an observer is presented with a stimulus with two targets of similar distance from the observer, one in a region with small projected size features and another in a region with large projected size features, the observer should overestimate the projected size of the target within the small-projected-size region. As an example stimulus, consider the one in the spot for 1A in Table 26.1 (this figure is the Ebbinghaus or Titchener illusion). The left side of the figure has, overall, smaller projected size features than the right side of the figure, and thus the left target, being probably nearer to the direction of motion, should undergo, in the next moment, a greater percentage increase in projected size. Because the two targets (i.e., the center circle on the left and the center circle on the right) have identical projected sizes, the left target will undergo, in the next moment, a greater increase in projected size than the one on the right, and perceiving-the-present expects observers to perceive the left target to project larger than the one on the right. Intuitively, the probable scene in 1A is of two identical circles on the left and right, at similar distance from the observer (see discussion below); but where the one on the left, being surrounded by smaller projected size features, is probably nearer to the observer's direction of motion and will undergo greater percentage growth in the next moment.

Consider as another example ecological regularity 1B, and the figure shown for it in Table 26.1. Here the target objects are objects moving at identical projected speed (indicated in the figure by arrows of identical length) over the horizontal lines, because (B) is the column for perceived projected speed. The horizontal lines on the left side of the stimulus have smaller projected size features (or greater projected spatial frequency), thus making that part of the stimulus probably nearer to the direction of motion (as indicated by (1)). We therefore expect that the left target will undergo a greater percentage increase in projected speed in the next moment, as the column heading states. Perceiving-the-present accordingly predicts that observers should perceive the projected speed of the target on the left to be greater than that of the same-speed target on the right (because that is how they would typically be in the next moment).

26.4.5 Arguments that the targets in illusions are treated as similarly distant from the observer

We will see later in this section that the illusions tend to be consistent with perceiving-the-present's predicted table of illusions in Table 26.1. However, unlike traditional explanations for illusions that rely upon claims about one target probably being farther away, our treatment supposes that the target objects tend to be treated by the visual system as if they are similarly distant. There are several reasons for believing that target objects in illusions and figures like those in Table 26.1 are treated by the visual system as similarly distant.

One reason is that the illusions do not change when strong cues are added that the targets are similarly distant. For example, in the Ponzo illusion in Fig. 26.1(b), the stimulus is ambiguous as to the distances of the two bars, and one possibility is that the stimulus is due to a scene where the top horizontal bar is farther from the observer (and another possibility is that the two targets are at similar distance from the observer). However, consider now the Orbison illusion in Fig. 26.1(b), where the two horizontal bars are now part of a square stimulus. It is highly probable that the square stimulus is due to a real-world square in the observer's frontoparallel plane, as opposed to a real-world trapezoid tilted backward in just the right manner to coincidentally project as a perfect square. In the Orbison illusion, then, the cues suggest that the upper and lower bars are probably at about the same distance from the observer (and this applies even more strongly for the grid illusion in Fig. 26.1(b)), and yet, importantly, the Ponzo-like illusion still is present.

The second reason for believing that target objects in illusions are treated by the visual system as similarly distant from the observer is that many illusions possess cues suggesting that the targets are, indeed, at similar distances from the observer. For example, in most illusions the two target stimuli are identical to each another (in projected size, shape, pattern, speed, and luminance), and the differences causing the illusion are in the surrounding stimuli, not in the targets themselves. This is true for most of the example stimuli in Table 26.1, namely 1A, 1B, 1D, 2A, 2B, 2C, 3A, 3B, 3C, 5A, 5B, 5C, 5D, 6A, 6B, 6C, 7A, 7B, 7C (where it is assumed that the vectors represent moving objects that are identical in projected size, shape, pattern, and luminance). This is, in fact, one of the central characteristics of a good illusion: that despite two stimuli being identical, they are perceived differently. But when two stimuli are identical, it significantly raises the probability that the targets are the same kind of object – it would be a rare coincidence that two different kinds of object in a scene would cause a stimulus with identical stimulus properties. It follows that if two objects are the same kind of object (having identical distal size), then because they project the same size in the stimulus, the two targets must probably be at the same distance from the observer. Thus, one of the central characteristics of a good illusion – having identical target stimuli – is itself a cue that the targets are at similar distance. Another kind of cue that two targets are at similar distance from the observer occurs in some illusions (e.g., in 2D, 3D, 4D, 6D, and 7D of Table 26.1), namely when the targets are the opposite ends of a single rectangularly projecting plane or grid (of uniform luminance). As mentioned in the previous paragraph, a rectangular stimulus is probably due to a real-world rectangle in the observer's frontoparallel plane – not due to a real-world trapezoid tilted in just such a manner as to coincidentally project rectangularly. Therefore, the opposite ends of the rectangular plane are at similar distances from the observer. Finally, illusions can have an even stronger cue that they are at similar distances from the observer, namely when identical binocular disparity is used (as in Row 4 of Table 26.1).

A third reason for believing that the visual system might treat the targets in illlusions as similarly distant from an observer is that, even if in a given stimulus the cues are weak that the targets are similarly distant, there are benefits for assuming, for perceiving-the-present compensation purposes, that the targets are at similar distances. For specificity, consider the Ponzo illusion (Fig. 26.1(b)), and consider two possible interpretations of the stimulus: (a) the lower bar is close to the observer but the upper horizontal bar is far, and (b) both bars are similarly close to the observer. Although in the previous paragraph we provided reason to believe that for the Ponzo the bars are probably at similar distance from the observer (because the two bar stimuli are identical), let us suppose now for the sake of example that these two possible interpretations are equally probable. If the observer perceives according to (a) but in fact (b) is the case, then the costs are potentially high because the upper bar will not be perceived veridically (it will in fact be moving closer and enlarging but will not be perceived to be so), and the observer is in danger of not interacting appropriately with the bars (e.g., collision). If, on the other hand, the observer perceives (b) but (a) is the case, then the costs are low because although the observer will not perceive the upper bar veridically (the upper bar's perceived angular size will be larger than the lower bar), the observer is far from the upper bar and not at risk of an inappropriate interaction with it (such as a collision).

26.4.6 Summary of perceiving-the-present prediction

In sum, Table 26.1 possesses two distinct but related kinds of content: (a) ecological regularities concerning how features tend to change in the next moment depending on their current features, and (b) predicted perceptions based on the perceiving-the-present hypothesis. Table 26.1 essentially predicts that there should be an underlying *pattern* to the kinds of illusions researchers have found, a pattern cutting across a broad spectrum of the visual perception literature. Rather than attempting to experimentally test each of these twenty-eight classes of illusion ourselves, here we have opted to test these predictions via a broad survey of the visual perception literature, which we discuss in the next subsection.

Before moving to the results of the meta-analysis, it is important to understand the kind of prediction we are making. Although for any predicted illusion class, illusions within that class may be empirically known, it is *not* empirically known what the *pattern* is across the twenty-eight distinct stimulus types. Our prediction concerns this underlying pattern: it should fit the regularities shown in Table 26.1. Note that this is significantly different from pooling together known cases of phenomena that appear to be consistent with one's hypothesis. As an analogy to the kind of prediction we are making, one of us (MAC) has a theory for how the number of neocortical areas should increase as a function of brain size (namely, a cube-root law). Given the prediction of the theory, MAC sought to test it by determining how the number of areas actually *does* increase as a function of brain size. To do this, he used the published literature to compile area counts and brain size measures and was able to find support for the prediction. But, one might complain, all the neocortical area and size information used in the plot were already known in the literature, and so no new prediction had been made by the theory. If we take such a complaint seriously, MAC would have to acquire area counts from new species before it could be called a prediction. The mistake in such a criticism is that, although each datum may have been known, the pattern made by compiling all the data had *not* been known. Similarly, the pattern across the twenty-eight distinct stimulus types was not known prior to our investigation; what the pattern is, then, is an open empirical question, and our theory predicts what the pattern's "shape" is. The following section empirically investigates this pattern.

Also, we must admit the limits of this predicted pattern. We cannot predict that *every* stimulus fitting within the stimulus type of one of the classes will have the predicted kind of illusion. The main reason for this is that the ecological regularities we presented earlier are statistical *tendencies*, not inviolate laws; they will sometimes fail, and some kinds of stimuli may be associated with those failure cases. This is again analogous to the neocortex-area example discussed above, where a predicted cube-root law relating number of areas to brain size does *not* mean that we do not expect outliers in some cases; of course we expect outliers, but we nevertheless expect the pattern to exist. As a cartoon example to make our point, imagine that observers typically steer away from spiders and steer toward lakes for the purpose of drinking. A stimulus with small projected size features on the left, but with spider shapes, and a large projected size on the right, with cues that it is a lake, may in fact be ecologically associated with movement toward the lake, with the larger projected size features; this would be counter to the ecological regularity in 1A of Table 26.1. We cannot possibly discount such possibilities. Our theory should be treated as a zeroth-order hypothesis. Nevertheless, our claim is that the *central* trend for any stimulus type will be as predicted by the general pattern. And, as we will see, there are indeed central trends in the literature for the kinds of illusions found of a given stimulus type. They are often the kinds of illusions the community has names for, because they are so strong or because they are so robust.

26.4.7 Testing the prediction in Table 26.1 via Table 26.2

Table 26.2 is like the "prediction table" Table 26.1, but it is the "illusion evidence table," where within each of its matrix boxes we have recorded illusions from the literature that fall within the corresponding stimulus class, along with citations. (A *stimulus class* consists of illusions where stimulus modality X affects perceived modality Y.) Any given stimulus from the literature falls within a well-defined stimulus class, and illusions were placed in the appropriate stimulus class of the table whether or not the illusion was as predicted by the hypothesis.

Although we can make no quantitative claim concerning the exhaustiveness of this literature search, Table 26.2 possesses more than 150 citations, the result of searching through on the order of 1000 papers. In a few illusion classes (3D, 4B, 4C, 6C, 6D, and 7C) we have found no existing research looking into the kind of stimuli about which the prediction concerns; in some of these cases the figure shown in Table 26.2 serves as a demonstration of the predicted phenomenon, but in other cases (the ones with question marks) it is left for future study. However, the powerful conclusion of the metareview is that, *for each of the remaining twenty-two illusion classes, the central phenomenon found in the literature accords with the prediction from Table 26.1.* That is, *the predicted pattern of illusions from Table 26.1 is largely confirmed by this study.* Furthermore, the theory allows the unification of more than 50 kinds of illusion into an orderly 7 by 4 table of illusion, demonstrating a broad empirical pattern across illusions not heretofore noticed. (For other observations of broad empirical patterns across illusions see Coren et al. [1976] and Gregory [2005].)

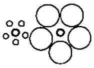
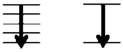
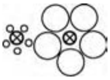
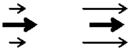
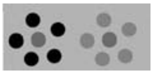
26.5 Conclusion

In this chapter we began by reviewing how the forward-motion perceiving-the-present theory explains many classical geometrical illusions. Our main goal here was to test two predictions of this theory.

The first prediction was that the converging lines found in classical geometrical illusion stimuli should be able to be replaced by dynamic optic flow (and with target stimuli that flash at some point in time rather than being always visible), and the illusions should still occur. We saw that this is the case in Section 26.3 where dynamic flow could be used to elicit analogs to the classical geometrical illusions, and we showed that the illusion could be modulated as expected by modulating the cued observer direction of motion, speed, and forward-versus-backward.

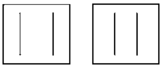
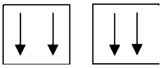
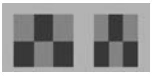
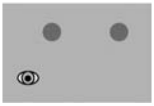
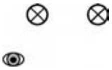
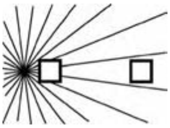
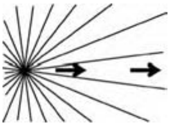
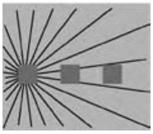
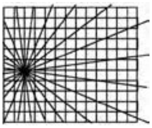
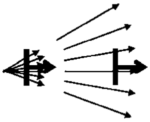
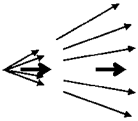
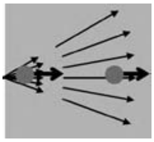
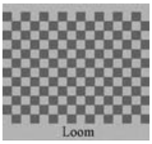
The second prediction concerned the observations that converging lines are just one of seven kinds of ecological cue to where the observer is heading, and that projected size is

Table 26.2 *7 × 4 table of 28 illusion classes catalogued from the visual perception literature. This table serves as data with which to test the prediction of Table 26.1*

	(A) . . . perceived (angular) *size*	**(B)** . . . perceived (angular) *speed*	**(C)** . . . perceived *luminance contrast*	**(D)** . . . perceived *distance*
(1) How (angular*) size* affects . . .	– **size contrast** (MacKay 1973; Klein et al. 1974; Georgeson 1980) – **Ebbinghaus/Titchener illusion** (Massaro & Anderson 1971; Coren & Girgus 1978; Weintraub & Schneck 1986) – **moon illusion** (Kaufman & Rock 1962; Rock & Kaufman 1962; Restle 1970; McCready 1986; Plug & Ross 1994; Kaufman & Kaufman 2000; Redding 2002) – **nearer horizon ⇒ larger** (Gilinsky 1955; Leibowitz & Harvey 1969) – **Oppel-Kundt/Botti illusion** (Oppel 1854; Lewis 1912; Rentschler et al. 1981; although counter at high spatial frequency, Rentschler et al. 1981) 	– **smaller background features ⇒ faster** (Brown 1931; Johansson 1950; Gogel & McNulty 1983) – **greater dot density of moving object ⇒ faster** (Watamaniuk et al., 1993) – **greater spatial frequency ⇒ faster** (Diener et al., 1976; McKee et al., 1986; although counter at high spatial frequency, Smith & Edgar 1990) – **smaller patterns ⇒ faster** (Snowden 1999) – **greater spatial frequency ⇒ less capture ⇒ greater relative speed** (De Valois & De Valois 1991) 	– **greater spatial frequency ⇒ greater "assimilation"** (Helson 1963; Steger 1969; Walker 1978) – **greater spatial frequency surround ⇒ lower contrast** (McCourt 1982; Yu et al. 2001) 	– **smaller surround ⇒ nearer** (Ebbinghaus/Titchener, McCready 1985 – **horizon moon appears nearer** (Boring 1943, 1962; Epstein et al. 1961; Rock & Kaufman 1962; McCready 1986) 
(2) How (angular) *speed* affects . . .	– **slower ⇒ larger** (Ansbacher 1944; Virsu et al. 1974; Parker 1981, 1983; Kaneko & Uchikawa 1993) – **longer presentation time ⇒ slower** (Katz et al. 1990; Treue et al. 1993) **⇒ lower spatial frequency** (Tynan & Sekuler 1974; Virsu et al. 1974; Virsu & Nyman 1974; Kulikowski 1975; Gelb & Wilson 1983; Maddess & Kulikowski 1999) 	– **motion contrast** (Loomis & Nakayama 1973; Walker & Powell 1974; Tynan & Sekuler 1975) 	– **lower speed surround ⇒ lower contrast** (Takeuchi & De Valois 2000) – **longer presentation time ⇒ slower** (Katz et al. 1990; Treue et al. 1993) **⇒ lower contrast** (Kulikowski 1972) 	– **lower speed surround ⇒ nearer** (motion parallax-induced depth contrast, Graham & Rogers 1982; Rogers & Graham 1983) 
(3) How *luminance contrast* affects . . .	– **greater contrast ⇒ stronger geometrical illusions** (Wallace 1975; Dworkin 1997) – **color equiluminance ⇒ absence of geometrical illusions** (Livingstone & Hubel 1987) – **greater contrast ⇒ larger** (Robinson 1954; Weale 1975; Georgeson 1980; De Weert et al. 1998) – **greater contrast ⇒ lower spatial frequency** (Virsu 1974; Virsu & Vuorinen 1975; Kulikowski 1975; Georgeson 1980; Gelb & Wilson 1983; Maddess & Kulikowski 1999) – **greater contrast ⇒ greater repulsion** (Rentschler et al. 1975) 	– **greater contrast ⇒ faster** (Hess 1904; Thompson 1982; Kooi et al. 1992; Stone & Thompson 1992; Müller & Greenlee 1994; Hawken et al. 1994; Ledgeway & Smith 1995; Gegenfurtner & Hawken 1996; Snowden et al. 1998; Blakemore & Snowden 1999; Brooks 2001) – **greater contrast moving surround ⇒ greater induced target speed** (Raymond & Darcangelo 1990) – **greater contrast ⇒ less capture ⇒ greater relative speed** (Ramachandran 1987; Ramachandran & Anstis 1990; Murakami & Shimojo 1993; Zhang et al. 1993) 	– **luminance contrast** (Ejima & Takahashi 1985; Chubb et al. 1989; Cannon & Fullenkamp 1991, 1993; Solomon et al. 1993; Snowden & Hammett 1998) 	– **color equiluminance ⇒ absence of depth** (Livingstone & Hubel 1987) – **greater contrast surround ⇒ nearer** (*BELOW*) 

(*cont.*)

Table 26.2 (*cont.*)

	(A) . . . perceived (angular) *size*	**(B)** . . . perceived (angular) *speed*	**(C)** . . . perceived *luminance contrast*	**(D)** . . . perceived *distance*
(4) How greater *distance* affects . . .	– **greater stereo depth ⇒ larger** (Enright 1989; Kaneko & Uchikawa 1997; De Weert et al. 1998) – **lower accommodation ⇒ larger** (Biersdorf et al. 1963) – **lower convergence ⇒ larger** (Thouless 1931; Heinemann et al. 1959; Biersdorf et al. 1963; Oyama & Iwawaki 1972; Komoda & Ono 1974) 	– **farther ⇒ faster** (follows from 4A) 	– **farther ⇒ lower contrast** (?) 	– **depth contrast** (Werner 1938; Harker 1962; Pastore 1964; Pastore & Terwilliger 1966; Anstis et al. 1978; Brookes & Stevens 1989; Pierce et al. 1998; van Ee et al. 1999; te Pas et al. 1997; Sato & Howard 2001) 
(5) How lower *eccentricity* affects . . .	– **nearer fovea ⇒ larger** (William James 1890/1950; von Helmholtz 1867/1962; Newsome 1972; Schneider et al. 1978; Georgeson 1980; Davis 1990) – **Hering illusion without the radial display** (von Helmholtz 1867/1962, see Liu & Schor 1998). – **foveal attraction or repulsion?** (Mateeff & Gourevich 1983; see Appendix for discussion) 	– **nearer fovea ⇒ faster** (Campbell & Maffei 1981; Tynan & Sekuler 1982; Pantle 1992; Schlykowa et al. 1993). – **nearer fovea ⇒ greater stereomotion speed** (Brooks & Mather 2000) – **nearer fovea ⇒ less capture ⇒ greater relative speed** (De Valois & De Valois 1991; Murakami & Shimojo 1993; Zhang et al. 1993). 	– **nearer fovea ⇒ lower contrast** (Georgeson 1991) 	– **nearer fovea ⇒ nearer** (barrel distortion of large frontoparallel planes, von Helmholtz 1867/1962) 
(6) How the *vanishing point* affects . . .	– **classical radial-display illusions such as Ponzo, Hering, Orbison** (Ehrenstein 1925; Orbison 1939; Berliner & Berliner 1948; Leibowitz et al. 1969; Coren et al. 1976; Coren & Girgus 1978; Schiffman & Thompson 1978; Weale 1978; Jordon & Randall 1987; Changizi 2001, 2003; Changizi & Widders 2002) 	– **nearer center of radial display ⇒ faster** (Swanston 1984; Cesàro & Agostini 1998) 	– **nearer center of radial ⇒ lower contrast** (*BELOW*) 	– **nearer center of radial display ⇒ nearer** (*BELOW*) 
(7) How the *focus of expansion* affects . . .	– **nearer FOE ⇒ larger** (Anstis 1989; Ramachandran & Anstis 1990; Whitaker et al. 1999) – **flow repulsion or capture** (Ramachandran 1987; Anstis 1989; Ramachandran & Anstis 1990; De Valois & De Valois 1991; Murakami & Shimojo 1993; Zhang et al. 1993; Whitney & Cavanagh 2000) – **forward induced motion ⇒ increased projected size** (Farnè 1972, 1977; Reinhardt-Rutland 1983; Wade & Swanston 1984; Whitaker et al. 1999) – **looming toward grid causes center to bow out** (Foster & Altschuler 2001) – **dynamic Zanker-Hering analog** (Zanker et al. 2001) 	– **nearer FOE ⇒ faster** (follows from 2B) – **looming illusions** (Widders color-balls in Changizi 2003) – **looming toward grid causes center to flow out faster** (Foster & Altschuler 2001) 	– **nearer FOE ⇒ lower contrast** (?) 	– **capture in depth** (Anstis 1989; Edwards & Badcock 2003) – **looming toward grid causes center to bulge** (Foster & Altschuler 2001) 

just one of four features that change in a predictable way as a function of distance in the visual field from the direction of motion. Therefore, the same forward-motion perceiving-the-present explanation for the classical geometrical illusions immediately makes 27 other predictions, namely a table of $7 \times 4 = 28$ predicted illusion classes, of which one holds the classical geometrical illusions, or a predicted pattern of illusion over 28 distinct stimulus types. We found that this pattern of illusions appears to exist via a metareview of the literature, amounting to a predictive success story for the forward-motion perceiving-the-present hypothesis. We note that the forward-motion perceiving-the-present theory was invented without knowing about this predicted pattern. Although the goal of the pattern found in Table 26.2 is to confirm a prediction of the forward-motion perceiving-the-present hypothesis (Table 26.1), we believe the empirical pattern is of significant interest for visual perception researchers, for it is the largest systematization of illusions thus far, it is theoretically motivated, and it unifies in an organized manner more than fifty kinds of illusion.

Appendix: Projected versus distal properties

Important for understanding our article is the distinction between the perception of distal properties and the perception of projected properties. *Distal properties* are features of the objects out there in the world. For example, the height of a tree (measured in meters) is a distal property, as is the surface reflectance (or lightness) of an object. *Projected properties*, on the other hand, concern only the nature of the light projected toward an observer at his/her particular location (not to be confused with retinal projection, although sometimes the retina may veridically record projected properties). For example, how much of the visual field is filled by a tree (measured in degrees, i.e., the angular size of the tree) is a projected property, as is the amount of light projected toward the eye (or brightness) from an object. It is useful to think of an imaginary projection sphere around a person's eyes, where the projection sphere lacks distance information. Projected properties are then properties measurable on this sphere. Projected properties are important to perceive (in addition to distal properties) because it is useful to perceive where things are in one's visual field (i.e., in which direction around oneself is an object), and once one is able to perceive where things are in one's visual field, the perception of projected size and speed follow because the former is just the projected distance (or visual angle) between two points in the visual field, and the latter is just the projected distance swept by a moving point during a unit time interval.

This "distal versus projected" distinction has been made often in the visual perception literature (Gibson 1950; Gilinsky 1955; Carlson 1960; Mack 1978; Rock 1983; Arend & Goldstein 1990; Sedgwick & Nicholis 1993; Gillam 1998; Palmer 1999; Changizi & Widders 2002), and perception of projected size (as opposed to distal size) has been observed a number of times over the history of visual perception (Reid 1813; Joynson 1949; Gibson 1950; Gilinsky 1955; Jenkin & Hyman 1959; Over 1960; Carlson 1960, 1962; Biersdorf et al. 1963; Rock & McDermott 1964; McCready 1965, 1985, 1986; Ono

1966; Baird 1968; Craig 1969; Leibowitz & Harvey 1969; Lucas 1969; Daniels 1972; Foley 1972; Angell 1974; Komoda & Ono 1974; Mack 1978; Sedgwick 1986; McKee & Welch 1989, 1992; Kaneko & Uchikawa 1993, 1997; Sedgwick & Nicholis 1993; Plug & Ross 1994; Gogel & Eby 1997). Researchers have also shown that observers make qualitatively different "size" judgments when given projected size instructions compared to when given distal size instructions (Gilinsky 1955; Jenkin & Hyman 1959; Carlson 1960, 1962; Biersdorf et al. 1963; Leibowitz & Harvey 1969): for stimuli with cues to the distal size, projected size instructions lead to judgments closely matching projected size, and distal size instructions lead to judgments closely matching distal size. Most of the literature on motion perception also recognizes the perception of projected properties, because perceived speed is nearly always measured in degrees/second. Furthermore, McKee and Welch (1989, 1992) provide evidence that discrimination for projected size and speed is often better, and never worse, than for distal size and speed. (Burbeck (1987) finds poorer proximal spatial frequency discriminations than distal, but see the discussion in McKee & Smallman (1998)). Measurements of perceived lightness (distal) versus perceived brightness (projected) have also been made (see Arend & Goldstein (1990)).

Acknowledgments

We wish to thank Zhi-Yong Yang, Timothy Hubbard, Andrew Wilson, David Eagleman, Bart Krekelberg, Bhavin Sheth, Daw-an Wu, Ladan Shams, Keith Rayner, Kyle Cave, and Patrick Cavanagh for their thoughtful comments and criticisms. Support was provided by the Sloan-Swartz Foundation and NIH grant 5F32EY015370–02.

References

Angell, R. B. (1974). The geometry of visibles. *Noûs* **8**: 87–117.

Ansbacher, H. L. (1944). Distortion in the perception of real movement. *J Exp Psychol* **34**: 1–23.

Anstis, S. (1989). Kinetic edges become displaced, segregated, and invisible. In D. M.-K. Lam & C. D. Gilbert (eds.), *Neural Mechanisms of Visual Perception* (247–260). The Woodlands, TX: Portfolia.

Anstis, S. M., Howard, I. P., & Rogers, B. (1978). A Craik-O'Brien-Cornsweet illusion for visual depth. *Vision Res* **18**: 213–217.

Arend, L. E., & Goldstein. R. (1990). Lightness and brightness over spatial illumination gradients. *J Opt Soc Am A* **7**: 1929–1936.

Baird, J. C. (1968). Toward a theory of frontal-size judgments. *Percept Psychophys* **4**: 49–53.

Berliner, A., & Berliner, S. (1948). The distortion of straight and curved lines in geometrical fields. *Am J Psychol* **61**: 153–166.

Berry, M. J. II, Brivanlou, I. H., Jordan, T. A., & Meister, M. (1999). Anticipation of moving stimuli by the retina. *Nature* **398**: 334–338.

Biersdorf, W. R., Ohwaki, S., & Kozil, D. J. (1963). The effect of instructions and oculomotor adjustments on apparent size. *Am J Psychol* **76**: 1–17.

Blakemore, M. R., & Snowden, R. J. (1999). The effect of contrast upon perceived speed: a general phenomenon? *Perception* **28**: 33–48.

Boring, E. G. (1943). The moon illusion. *Am J Phys* **11**: 55–60.

Boring, E. G. (1962). On the moon illusion – a letter. *Science* **137**: 902–906.

Brookes, A., & Stevens, K. A. (1989). The analogy between stereo depth and brightness. *Perception* **18**: 601–614.

Brooks, K. (2001). Stereomotion speed perception is contrast dependent. *Perception* **30**: 725–731.

Brooks, K., & Mather, G. (2000). Perceived speed of motion in depth is reduced in the periphery. *Vision Res* **40**: 3507–3516.

Brown, J. F. (1931). The visual perception of velocity. *Psychologische Forschung* **14**: 199–232.

Burbeck, C. A. (1987). Locus of spatial-frequency discrimination. *J Opt Soc Am A* **4**: 1807–1813.

Burr, D. (2000). Motion vision: are "speed lines" used in human visual motion? *Curr Biol* **10**: R440–R443.

Burr, D. C., & Ross, J. (2002). Direct evidence that "speedlines" influence motion mechanisms. *J Neurosci* **22**: 8661–8664.

Campbell, F. W., & Maffei, L. (1981). The influence of spatial frequency and contrast on the perception of moving patterns. *Vision Res* **21**: 713–721.

Cannon, M. W., & Fullenkamp, S. C. (1991). Spatial interactions in apparent contrast: inhibitory effects among grating patterns of different spatial frequencies, spatial positions and orientations. *Vision Res* **31**: 1985–1998.

Cannon, M. W. Jr., & Fullenkamp, S. C. (1993). Spatial interactions in apparent contrast: individual differences in enhancement and suppression effects. *Vision Res* **33**: 1685–1695.

Carlson, V. R. (1960). Overestimation in size-constancy judgments. *Am J Psychol* **73**: 199–213.

Carlson, V. R. (1962). Size-constancy judgments and perceptual compromise. *J Exp Psychol* **63**: 68–73.

Cesàro, A. L., & Agostini, T. (1998). The trajectory of a dot crossing a pattern of tilted lines is misperceived. *Percept Psychophys* **60**: 518–523.

Changizi, M. A. (2001). "Perceiving the present" as a framework for ecological explanations of the misperception of projected angle and angular size. *Perception* **30**: 195–208.

Changizi, M. A. (2003). *The Brain from 25,000 Feet: High Level Explorations of Brain Complexity, Perception, Induction and Vagueness*. Dordrecht: Kluwer Academic.

Changizi, M. A. (2009). *The Vision Revolution*. Dallas, TX: Benbella.

Changizi, M. A., & Widders, D. (2002). Latency correction explains the classical geometrical illusions. *Perception* **31**: 1241–1262.

Chubb, C., Sperling, G., & Solomon, J. A. (1989). Texture interactions determine perceived contrast. *Proc Natl Acad Sci USA* **86**: 9631–9635.

Coren, S., & Girgus, J. S. (1978). *Seeing Is Deceiving: The Psychology of Visual Illusions*. Hillsdale, NJ: Erlbaum.

Coren, S., Girgus, J. S., Erlichman, H., & Hakstian, A. R. (1976). An empirical taxonomy of visual illusions. *Percept Psychophys* **20**: 129–137.

Craig, E. J. (1969). Phenomenal geometry. *Br J Philos Sci* **20**: 121–134.

Cutting, J. E. (2002). Representing motion in a static image: constraints and parallels in art, science, and popular culture. *Perception* **31**: 1165–1194.

Cutting, J. E., & Vishton, P. M. (1995). Perceiving layout and knowing distance: the integration, relative potency, and contextual use of different information about depth. In W. Epstein & S. Rogers (eds.), *Handbook of Perception and Cognition. Vol. 5: Perception of Space and Motion* (69–117). San Diego: Academic Press.

Daniels, N. (1972). Thomas Reid's discovery of a non-euclidean geometry. *Philosophy of Science* **39**: 219–234.

Davis, E. T. (1990). Modeling shifts in perceived spatial frequency between the fovea and periphery. *J Opt Soc Am A* **7**: 286–296.

De Valois, R. L., & De Valois, K. K. (1991). Vernier acuity with stationary moving gabors. *Vision Res* **31**: 1619–1626.

De Weert, C. M. M., Snoeren, P. R., & Puts, M. J. H. (1998). Mutual dependence of luminance, size, and disparity in depth, size and luminance discrimination tasks. *Perception* **27** (suppl.): 111.

Diener, H. C., Wist, E. R., Dichgans, J., & Brandt, T. (1976). The spatial frequency effect on perceived velocity. *Vision Res* **16**: 169–176.

Dworkin, L. (1997). The effects of brightness contrast and stimulus presentation and duration on the magnitude of the Oppel-Kundt illusion. Unpublished thesis, Department of Psychology, Concordia University.

Edwards, M., & Badcock, D. R. (2003). Motion distorts perceived depth. *Vision Res* **43**: 1799–1804.

Ehrenstein, W. (1925). Versuche ueber die Beziehungen zwischen Bewegungs- und Gestaltwahrnehmung. (Experiments on the relationships between the perception of motion and of gestalt). *Zeitschrift fuer Psychologie* **96**: 305–352.

Ejima, Y., & Takahashi, S. (1985). Apparent contrast of a sinusoidal grating in the simultaneous presence of peripheral gratings. *Vision Res* **25**: 1223–1232.

Enright, J. T. (1989). Manipulating stereopsis and vergence in an outdoor setting: moon, sky and horizon. *Vision Res* **29**: 1815–1824.

Epstein, W., Park, J., & Casey, A. (1961). The current status of the size-distance hypothesis. *Psychol Bull* **58**: 491–514.

Farnè, M. (1972). Studies on induced motion in the third dimension. *Perception* **1**: 351–357.

Farnè, M. (1977). Motion in depth induced by brightness changes in the background. *Perception* **6**: 295–297.

Foley, J. M. (1972). The size-distance relation and intrinsic geometry of visual space: implications for processing. *Vision Res* **12**: 323–332.

Foster, C., & Altschuler, E. L. (2001). The bulging grid. *Perception* **30**: 393–395.

Gegenfurtner, K. R., & Hawken, M. J. (1996). Perceived velocity of luminance, chromatic and non-fourier stimuli: influence of contrast and temporal frequency. *Vision Res* **36**: 1281–1290.

Geisler, W. S. (1999). Motion streaks provide a spatial code for motion direction. *Nature* **400**: 65–69.

Geisler, W. S., Albrecht, D. G., Crane, A. M., & Stern, L. (2001). Motion direction signals in the primary visual cortex of cat and monkey. *Vis Neurosci* **18**: 501–516.

Gelb, D. J., & Wilson, H. R. (1983). Shifts in perceived size as a function of contrast and temporal modulation. *Vision Res* **23**: 71–82.

Georgeson, M. A. (1980). Spatial frequency analysis in early visual processing. *Philos Trans R Soc Lond B Biol Sci* **290**: 11–22.

Georgeson, M. A. (1991). Contrast overconstancy. *J Opt Soc Am A* **8**: 579–586.

Gibson, J. J. (1950). *The Perception of the Visual World.* Boston: Houghton Mifflin.

Gilinsky, A. S. (1955). The effect of attitude upon the perception of size. *Am J Psychol* **68**: 173–192.

Gillam, B. J. (1998). Illusions at century's end. In J. Hochberg (ed.), *Perception and Cognition at Century's End* (98–137). San Diego: Academic Press.

Gogel, W. C., & Eby, D. W. (1997). Measures of perceived linear size, sagittal motion, and visual angle from optical expansions and contractions. *Percept Psychophys* **59**: 783–806.

Gogel, W. C., & McNulty, P. (1983). Perceived velocity as a function of reference mark density. *Scand J Psychol* **24**: 257–265.

Gregory, R. (2005). The Medawar Lecture 2001: knowledge for vision, vision for knowledge. *Philos Trans R Soc Lond B Biol Sci* **360**: 1231–1251.

Harker, G. S. (1962). Apparent frontoparallel plane, stereoscopic correspondence, and induced cyclotorsion of the eyes. *Percept Mot Skills* **14**: 75–87.

Hawken, M. J., Gegenfurtner, K. R., & Tang, C. (1994). Contrast dependence of colour and luminance motion mechanisms in human vision. *Nature* **367**: 268–270.

Heinemann, E. G., Tulving, E., & Nachmias, J. (1959). The effect of oculomotor adjustments on apparent size. *Am J Psychol* **72**: 32–45.

Helson, H. (1963). Studies of anomalous contrast and assimilation. *J Opt Soc Am* **53**: 179–184.

Hess, C. V. (1904). Untersuchungen über den Erregungsvorgang in Sehorgan bei Kurz- und bei länger dauernder Reizung. *Pflügers Arch Gesamte Physiol* **101**: 226–262.

James, W. (1950). *Principles of Psychology, Vol. II* (New York: Dover), originally published 1890.

Jenkin, N., & Hyman, R. (1959). Attitude and distance-estimation as variables in size-matching. *Am J Psychol* **72**: 68–76.

Johansson, G. (1950). Configurations in the perception of velocity. *Acta Psychologica* **7**: 25–79.

Jordan, K., & Randall, J. (1987). The effects of framing ratio and oblique length on Ponzo illusion magnitude. *Percept Psychophys* **41**: 435–439.

Joynson, R. B. (1949). The problem of size and distance. *Q J Exp Psychol* **1**: 119–135.

Kaneko, H., & Uchikawa, K. (1993). Apparent relative size and depth of moving objects. *Perception* **22**: 537–547.

Kaneko, H., & Uchikawa, K. (1997). Perceived angular and linear size: the role of binocular disparity and visual surround. *Perception* **26**: 17–27.

Katz, E., Gizzi, M. S., Cohen, B., & Malach, R. (1990). The perceived speed of object motion varies inversely with distance travelled. *Perception* **19**: 387.

Kaufman, L., & Kaufman, J. H. (2000). Explaining the moon illusion. *Proc Natl Acad Sci USA* **97**: 500–505.

Kaufman, L., & Rock, I. (1962). The moon illusion, I. *Science* **136**: 953–961.

Khurana, B., Watanabe, R., & Nijhawan, R. (2000). The role of attention in motion extrapolation: are moving objects "corrected" or flashed objects attentionally delayed? *Perception* **29**: 675–692.

Klein, S., Stromeyer, C. F. III, & Ganz, L. (1974). The simultaneous spatial frequency shift: a dissociation between the detection and perception of gratings. *Vision Res* **14**: 1421–1432.

Komoda, M. K., & Ono, H. (1974). Oculomotor adjustments and size-distance perception. *Percept Psychophys* **15**: 353–360.

Kooi, F. L., De Valois, K. K., Grosof, D. H., & De Valois, R. L. (1992). Properties of the recombination of one-dimensional motion signals into a pattern motion signal. *Percept Psychophys* **52**: 415–424.

Kourtzi, Z., & Kanwisher, N. (2000). Activation in human MT/MST by static images with implied motion. *J Cogn Neurosci* **12**: 48–55.

Krekelberg, B., Dannenberg, S., Hoffmann, K-P., Bremmer, F., & Ross, J. (2003). Neural correlates of implied motion. *Nature* **424**: 674–677.

Krekelberg, B., Vatakis, A., & Kourtzi, Z. (2005). Implied motion from form in the human visual cortex. *J Neurophysiol* **94**: 4373–4386.

Kulikowski, J. J. (1972). Relation of psychophysics to electrophysiology. *Trace, Paris* **6**: 64–69.

Kulikowski, J. J. (1975). Apparent fineness of briefly presented gratings: balance between movement and pattern channels. *Vision Res* **15**: 673–680.

Ledgeway, T., & Smith, A. T. (1995). The perceived speed of second-order motion and its dependence on stimulus contrast. *Vision Res* **35**: 1421–1434.

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

Leibowitz, H. W., & Harvey, L. O., Jr. (1969). Effect of instructions, environment, and type of test object on matched size. *J Exp Psychol* **81**: 36–43.

Lennie, P. (1981). The physiological basis of variations in visual latency. *Vision Res* **21**: 815–824.

Lewis, C. F., & McBeath, M. K. (2004). Bias to experience approaching motion in a three-dimensional virtual environment. *Perception* **33**: 259–276.

Lewis, E. O. (1912–1913). The illusion of filled and unfilled space. *Br J Psychol* **5**: 36–50.

Liu, L., & Schor, C. M. (1998). Functional division of the retina and binocular correspondence. *J Opt Soc Am A* **15**: 1740–1755.

Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *J Neurosci* **7**: 3416–3468.

Loomis, J. M., & Nakayama, K. (1973). A velocity analogue of brightness contrast. *Perception* **2**: 425–428.

Lucas, J. R. (1969). Euclides ab omni naevo vindicatus. *Br J Philos Sci* **20**: 1–11.

Mack, A. (1978). Three modes of visual perception. In M. H. Pick & E. Saltzman (eds.), *Modes of Perceiving and Information Processing* (171–186). Hillsdale, NJ: Erlbaum.

MacKay, D. M. (1973). Lateral interaction between neural channels sensitive to texture density? *Nature* **245**: 159–161.

Maddess, T., & Kulikowski, J. J. (1999). Apparent fineness of stationary compound gratings. *Vision Res* **39**: 3404–3416.

Massaro, D. W., & Anderson, N. H. (1971). Judgmental model of the Ebbinghaus illusion. *J Exp Psychol* **89**: 147–151.

Mateeff, S., & Gourevich, A. (1983). Peripheral vision and perceived visual direction. *Biol Cybern* **49**: 111–118.

Maunsell, J. H. R., & Gibson, J. R. (1992). Visual response latencies in striate cortex of the macaque monkey. *J Neurophysiol* **68**: 1332–1344.

McBeath, M. K., Morikawa, K., & Kaiser, M. K. (1992). Perceptual bias for forward-facing motion. *Psychol Sci* **3**: 362–367.

McCourt, M. E. (1982). A spatial frequency dependent grating-induction effect. *Vision Res* **22**: 119–134.

McCready, D. (1965). Size-distance perception and accommodation-convergence micropsia – a critique. *Vision Res* **5**: 189–206.

McCready, D. (1985). On size, distance, and visual angle perception. *Percept Psychophys* **37**: 323–334.

McCready, D. (1986). Moon illusions redescribed. *Percept Psychophys* **39**: 64–72.

McKee, S. P., Silverman, G. H., & Nakayama, K. (1986). Precise velocity discrimination despite variations in temporal frequency and contrast. *Vision Res* **26**: 609–619.

McKee, S. P., & Smallman, H. S. (1998). Size and speed constancy. In V. Walsh & J. Kulikowski (eds.), *Perceptual Constancy: Why Things Look As They Do* (373–408). Cambridge: Cambridge University Press.

McKee, S. P., & Welch, L. (1989). Is there a constancy for velocity? *Vision Res* **29**: 553–561.

McKee, S. P., & Welch, L. (1992). The precision of size constancy. *Vision Res* **32**: 1447–1460.

Müller, R., & Greenlee, M. W. (1994). Effect of contrast and adaptation on the perception of the direction and speed of drifting gratings. *Vision Res* **34**: 2071–2092.

Murakami, I., & Shimojo, S. (1993). Motion capture changes to induced motion at higher luminance contrasts, smaller eccentricities, and larger inducer sizes. *Vision Res* **33**: 2091–2107.

Newsome, L. R. (1972). Visual angle and apparent size of objects in peripheral vision. *Percept Psychophys* **12**: 300–304.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature* **370**: 256–257.

Nijhawan, R. (1997). Visual decomposition of colour through motion extrapolation. *Nature* **386**: 66–69.

Nijhawan, R. (2001). The flash-lag phenomenon: object motion and eye movements. *Perception* **30**: 263–282.

Nijhawan, R. (2002). Neural delays, visual motion and the flash-lag effect. *Trends Cogn Sci* **6**: 387–393.

Ono, H. (1966). Distal and projected size under reduced and non-reduced viewing conditions. *Am J Psychol* **79**: 234–241.

Oppel, J. J. (1854–1855). *Über geometrisch-optische Tauschungen. Jahresbericht des Frankfurter Vereins* (37–47).

Orbison, W. D. (1939). Shape as a function of the vector-field. *Am J Psychol* **52**: 31–45.

Over, R. (1960). The effect of instructions on size-judgments under reduction-conditions. *Am J Psychol* **73**: 599–602.

Oyama, T., & Iwawaki, S. (1972). Role of convergence and binocular disparity in size constancy. *Psychol Forsch* **35**: 117–130.

Palmer, S. E. (1999). *Vision Science: Photons to Phenomenology*. Cambridge, MA: MIT Press.

Pantle, A. (1992). Immobility of some second-order stimuli in human peripheral vision. *J Opt Soc Am A* **9**: 863–867.

Parker, A. (1981). Shifts in perceived periodicity induced by temporal modulation and their influence on the spatial frequency tuning of two aftereffects. *Vision Res* **21**: 1739–1747.

Parker, A. (1983). The effects of temporal modulation on the perceived spatial structure of sine-wave gratings. *Perception* **12**: 663–682.

Pastore, N. (1964). Induction of a stereoscopic depth effect. *Science* **144**: 888.

Pastore, N., & Terwilliger, M. (1966). Induction of stereoscopic depth effects. *Br J Psychol* **57**: 201–202.

Pavlova, M., Krägeloh-Mann, I., Birbaumer, N., & Sokolov, A. (2002). Biological motion shown backwards: the apparent-facing effect. *Perception* **31**: 435–443.

Pierce, B. J., Howard, I. P., & Feresin, C. (1998). Depth interactions between inclined and slanted surfaces in vertical and horizontal orientations. *Perception* **27**: 87–103.

Plug, C., & Ross, H. E. (1994). The natural moon illusion: a multifactor angular account. *Perception* **23**: 321–333.

Ramachandran, V. S. (1987). Interaction between colour and motion in human vision. *Nature* **328**: 645–647.

Ramachandran, V. S., & Anstis, S. M. (1990). Illusory displacement of equiluminous kinetic edges. *Perception* **19**: 611–616.

Raymond, J. E., & Darcangelo, S. M. (1990). The effect of local luminance contrast on induced motion. *Vision Res* **30**: 751–756.

Redding, G. M. (2002). A test of size-scaling and relative-size hypotheses for the moon illusion. *Percept Psychophys* **64**: 1281–1289.

Regan, D., & Beverly, K. I. (1982). How do we avoid confounding the direction we are looking and the direction we are going? *Science* **215**: 194–196.

Reid, T. (1813). Works. In Four Volumes. Stewart, Dugald, et al. *Inquiry into the Human Mind*. Charlestown: Samuel Etheridge, Jr., in vol 1.

Reinhardt-Rutland, A. H. (1983). Induced movement-in-depth: Relative location of static stimulus and direction asymmetry. *Percept Mot Skills* **57**: 255–258.

Rentschler, I., Hilz, R., & Grimm, W. (1975). Processing of positional information in the human visual system. *Nature* **253**: 444–445.

Rentschler, I., Hilz, R., Sütterlin, C., & Noguchi, K. (1981). Illusions of filled lateral and angular extent. *Exp Brain Res* **44**: 154–158.

Restle, F. (1970). Moon illusion explained on the basis of relative size. *Science* **167**: 1092–1096.

Robinson, E. J. (1954). The influence of photometric brightness on judgments of size. *Am J Psychol* **67**: 464–474.

Rock, I. (1983). *The Logic of Perception*. Cambridge, MA: MIT Press.

Rock, I., & Kaufman, L. (1962). The moon illusion, II. *Science* **136**: 1023–1031.

Rock, I., & McDermott, W. (1964). The perception of visual angle. *Acta Psychologica* **22**: 119–134.

Rogers, B. J., & Graham, M. E. (1983). Anisotropies in the perception of three-dimensional surfaces. *Science* **221**: 1409–1411.

Ross, J., Badcock, D. R., & Hayes, A. (2000). Coherent global motion in the absence of coherent velocity signals. *Curr Biol* **10**: 679–682.

Sato, M., & Howard, I. P. (2001). Effects of disparity-perspective cue conflict on depth contrast. *Vision Res* **41**: 415–426.

Schiffman, H. R., & Thompson, J. G. (1978). The role of apparent depth and context in the perception of the Ponzo illusion. *Perception* **7**: 47–50.

Schlag, J., Cai, R. H., Dorfman, A., Mohempour, A., & Schlag-Rey, M. (2000). Extrapolating movement without retinal motion. *Nature* **403**: 38–39.

Schlykowa, L., Ehrenstein, W. H., Cavonius, C. R., & Arnold, B. E. (1993). *Perception* 22 (suppl.): 97.

Schmolesky, M. T., Wang, Y., Hanes, D. P., Thompson, K. G., Leutger, S., Schall, J. D., et al. (1998). Signal timing across the macaque visual system. *J Neurophysiol* **79**: 3272–3278.

Schneider, B., Ehrlich, D. J., Stein, R., Flaum, M., & Mangel, S. (1978). Changes in the apparent lengths of lines as a function of degree of retinal eccentricity. *Perception* **7**: 215–223.

Sedgwick, H. A. (1986). In K. R. Boff, L. Kaufman, & J. P. Thomas (eds.), *Handbook of Perception and Human Performance, Vol. 1: Sensory Processes and Perception* (21.1–21.57). New York: Wiley.

Sedgwick, H. A., & Nicholis, A. L. (1993). Interaction between surface and depth in the Ponzo illusion. *Invest Ophthalmol Vis Sci* **34**: 1184.

Sheth, B. R., Nijhawan, R., & Shimojo, S. (2000). Changing objects lead briefly flashed ones. *Nat Neurosci* **3**: 489–495.

Smith, A. T., & Edgar, G. K. (1990). The influence of spatial frequency on perceived temporal frequency and perceived speed. *Vision Res* **30**: 1467–1474.

Snowden, R. J. (1999). The bigger they are the slower they move: the effects of field size on speed discrimination. *Perception* **28**: 24S.

Snowden, R. J., & Hammett, S. T. (1998). The effects of surround contrast on contrast thresholds, perceived contrast and contrast discrimination. *Vision Res* **38**: 1935–1945.

Snowden, R. J., Stimpson, N., & Ruddle, R. A. (1998). Speed perception fogs up as visibility drops. *Nature* **392**: 450.

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 Res* **33**: 2671–2683.

Steger, J. A. (1969). Visual lightness assimilation and contrast as a function of differential stimulation. *Am J Psychol* **82**: 56–72.

Stone, L. S., & Thompson, P. (1992). Human speed perception is contrast dependent. *Vision Res* **32**: 1535–1549.

Swanston, M. T. (1984). Displacement of the path of perceived movement by intersection with static contours. *Percept Psychophys* **36**: 324–328.

Takeuchi, T., & De Valois, K. K. (2000). Modulation of perceived contrast by a moving surround. *Vision Res* **40**: 2697–2709.

te Pas, S. F., Rogers, B. J., & Ledgeway, T. (1997). *A curvature contrast effect for stereoscopically-defined surfaces*. Applied Vision Association Meeting on Depth Perception, United Kingdom.

Thompson, P. (1982). Perceived rate of movement depends on contrast. *Vision Res* **22**: 377–380.

Thouless, R. H. (1931). Phenomenal regression to the real object. *I. Br J Psychol* **21**: 339–359.

Tinbergen, N. (1939). Why do the birds behave the way they do? *Bird Lore* **41**: 23–30.

Tinbergen, N. (1951). *The Study of Instinct.* Oxford: Oxford University Press.

Treue, S., Snowden, R. J., & Andersen, R. A. (1993). The effect of transiency on perceived velocity of visual patterns: a case of "temporal capture." *Vision Res* **33**: 791–798.

Tynan, P., & Sekuler, R. (1974). Perceived spatial frequency varies with stimulus duration. *J Opt Soc Am* **64**: 1251–1255.

Tynan, P., & Sekuler, R. (1975). Simultaneous motion contrast: velocity, sensitivity and depth response. *Vision Res* **15**: 1231–1238.

Tynan, P. D., & Sekuler, R. (1982). Motion processing in peripheral vision: reaction time and perceived velocity. *Vision Res* **22**: 61–68.

van Ee, R., Banks, M. S., & Backus, B. T. (1999). An analysis of binocular slant contrast. *Perception* **28**: 1121–1145.

Virsu, V. (1974). Dark adaptation shifts apparent spatial frequency. *Vision Res* **14**: 433–435.

Virsu, V., & Nyman, G. (1974). Monophasic temporal modulation increases apparent spatial frequency. *Perception* **3**: 337–363.

Virsu, V., Nyman, F., & Lehtiö, P. K. (1974). Diphasic and polyphasic temporal modulations multiply apparent spatial frequency. *Perception* **3**: 323–336.

Virsu, V., & Vuorinen, R. (1975). Dark adaptation and short-wavelength backgrounds decrease perceived size. *Perception* **4**: 19–34.

von Helmholtz, H. (1867/1962). *Treatise on Physiological Optics vol. 3* (New York: Dover, 1962); English translation by JPC Southall for the Optical Society of America (1925) from the 3rd German edition of *Hundbuch der Physiologischen Optik* (Voss, Hamburg, 1910; first published in 1867, Voss, Leipzig).

Wade, N. J., & Swanston, M. T. (1984). Illusions of size change in dynamic displays. *Percept Psychophys* **35**: 286–290.

Walker, P., & Powell, D. J. (1974). Lateral interaction between neural channels sensitive to velocity in the human visual system. *Nature* **252**: 732–733.

Wallace, G. K. (1975). The effect of contrast on the Zollner illusion. *Vision Res* **15**: 963–966.

Wann, J. P., & Swapp, D. K. (2000). Why you should look where you are going. *Nat Neurosci* **3**: 647–648.

Watamaniuk, S. N. J., Grzywacz, N. M., & Yuille, A. L. (1993). Dependence of speed and direction perception on cinematogram dot density. *Vision Res* **33**: 849–859.

Weale, R. A. (1975). Apparent size and contrast. *Vision Res* **15**: 949–955.

Weale, R. A. (1978). Experiments on the Zöllner and related optical illusions. *Vision Res* **18**: 203–208.

Weintraub, D. J., & Schneck, M. K. (1986). Fragments of Delboeuf and Ebbinghaus illusions: contour/context explorations of misjudged circle size. *Percept Psychophys* **40**: 147–158.

Werner, H. (1938). Binocular depth contrast and the conditions of the binocular field. *Am J Psychol* **51**: 489–497.

Whitaker, D., McGraw, P. V., & Pearson, S. (1999). Non-veridical size perception of expanding and contracting objects. *Vision Res* **39**: 2999–3009.

Whitney, D., & Cavanagh, P. (2000). Motion distorts visual space: shifting the perceived position of remote stationary objects. *Nat Neurosci* **3**: 954–959.

Wilkie, R. M., & Wann, J. P. (2003). Eye-movements aid the control of locomotion. *J Vis* **3**: 677–684.

Yu, C., Klein, S. A., & Levi, D. M. (2001). Surround modulation of perceived contrast and the role of brightness induction. *J Vis* **1**: 18–31.

Zanker, J. M., Quenzer, T., & Fahle, M. (2001). Perceptual deformation induced by visual motion. *Naturewissenschaften* **88**: 129–132.

Zhang, J., Yeh, S-L., & De Valois. (1993). Motion contrast and motion integration. *Vision Res* **33**: 2721–2732.

27

History and theory of flash-lag: past, present, and future

GERRIT W. MAUS, BEENA KHURANA, AND ROMI NIJHAWAN

Summary

Some basic versions of the flash-lag effect have been known since the early decades of the twentieth century. Intriguingly, neural delays were as central in the early attempts at explaining the effect, as they are in the more recent investigations into its cause. For a changing visual stimulus a delayed registration of the stimulus by the central nervous system (CNS) constitutes an "error" between the instantaneously registered state of the stimulus on the one hand and its physical state on the other. Therefore, for animals to acquire food, mate, and avoid predators, compensation of sensory delays is essential. One may ask which component(s) of the CNS compensate for visual delays. Logically compensation could be carried out either by visual or motor mechanisms, or both. The motion extrapolation account of the flash-lag effect challenged the dominant view that only motor mechanisms compensate for visual delays, suggesting instead that visual mechanisms also contribute. Controversy fueled by empirical observations with unpredictable motion, in particular the flash-initiated and flash-terminated conditions of the flash-lag effect, soon followed; prima facie motion extrapolation could not accommodate these results. Armed with these challenging findings (primarily) several alternative accounts of flash-lag were proposed. In light of new developments, this chapter evaluates the motion extrapolation, motion sampling, motion integration, postdiction, differential latency, and attentional cuing accounts of flash-lag.

27.1 Introduction: time delays in the nervous system and need for compensation

One key function of neural systems is to detect change in external environments or internal states. A fundamental type of environmental change is movement of a stimulus across a receptor surface. Movement, for example on the skin of remote body parts or in the periphery of the retina, attracts attention such that organisms direct high-resolution sensory surfaces (e.g., fingertips or the fovea) to the locations of movement (James 1890/1952). Movement often stems from another organism, which could either be a conspecific, a potential mate or competitor, or another species, a potential prey or predator. Even inanimate moving objects are likely to be of importance, as they could present an obstacle or possible threat while the animal itself is in motion. In all cases, having information about the exact position of anything that moves is a prerequisite for successful interaction with the moving object, whether the goal be to avoid or intercept.

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

There are several problems the nervous system has to solve before it can determine the position of a moving object. One general property of biological neural systems that compounds the localization of moving objects is that biological processes take time. Processes such as phototransduction in receptor cells, signal conductance along nerve fibers, and synaptic transmission all rely on electrochemical mechanisms that work on the order of several milliseconds. The contribution from each of these processes can add up to behaviorally significant time delays. Problematically, during the time delay the environment is dynamic; other organisms and objects move and change position. In principle, it is impossible for an organism to have absolute knowledge about the positions of all relevant objects. However, to be able to interact with moving objects in a dynamic environment, organisms have developed sophisticated mechanisms to overcome these processing delays.

Without any kind of compensation for delays in neural processing it would be impossible for an organism to successfully interact with moving objects, as the position represented at any stage in the central nervous system would always lag behind the physical position of the object in the environment. Hence action directed at this lagging position would fail to make contact with the object in question. Wild cats would jump short of fast-moving prey and forest-dwelling animals would run into trees. Modern high-speed ball games provide nice examples to evaluate the potential error of localizations without compensation. If we assume neural delays from the retina to some cortical processing stage of about 100 msec, a tennis ball flying at a speed of 100 km/hr (27.8 msec^{-1}) would be misrepresented by 2.78 m. Obviously, in many instances predators are able to catch fast-moving prey and players are able to hit tennis balls. Therefore, compensation of delays at some stage of the neural processing is mandatory.

Given the sluggishness of muscle responses and the inertia in the actual movement of limbs, compensation for delays in the motor system is well established in neuroscience (Ghez & Krakauer 2000). It is commonly accepted that delays in other (sensory) parts of the nervous system are also compensated for at the motor stage of neural processing (Jordan 1995). However, there is no a priori reason why all compensation should take place at one summative stage, as processing delays are prevalent in all parts of the nervous system. Visual effects involving moving objects have led to the proposal of predictive mechanisms that compensate for delays in sensory pathways, thereby facilitating the accurate perception of positions (Nijhawan 1994). This discussion of neural delays and compensatory mechanisms has been revitalized by the flash-lag effect (Nijhawan 1994) in which a brief flash, presented in spatial alignment with a moving object, appears to lag behind the moving object (Fig. 27.1).

27.2 History of flash-lag

Research in the last fifteen years, following the publication of Nijhawan's (1994) motion extrapolation proposal, has engendered considerable debate in the scientific community on the underlying mechanisms of the flash-lag effect. However, around the time of the birth of experimental psychology as an independent field, similar perceptual phenomena

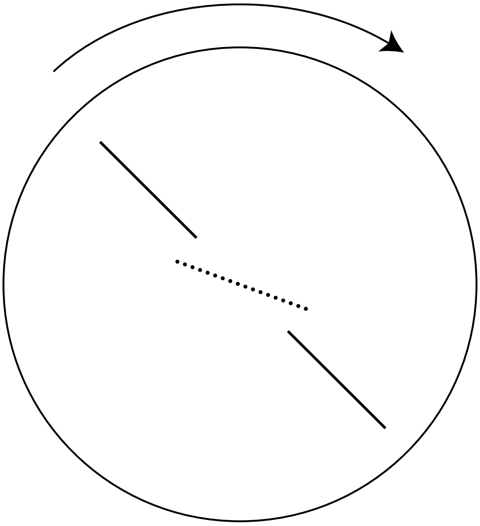


Fig. 27.1 In a dark room a single physical rod, made of three segments, rotates in the direction of the arrow. The two outer segments (solid lines) are illuminated with a continuous light source, whereas the central segment (dashed line) is illuminated with a briefly flashed light source. The figure shows the percept of the observers, with the central segment lagging behind (from Nijhawan 1992; see Appendix).

relating to the localization of moving objects had been studied. For example, in astronomical measurements, it was often necessary to record the accurate transit times of stars crossing the North-South meridian. To measure these time points astronomers had to record the moment when a star moved across a meridian wire in a telescopic view by estimating the perceived position of the star at certain clock beats (Mollon & Perkins 1996). This method relied on the perception of the instantaneous visual position of a moving object at the instant an auditory time marker sounded, thus potentially giving rise to a cross-modal flash-lag effect (Alais & Burr 2003). Observations of intersubjective differences in this method triggered the first experimental investigations of reaction times and temporal order judgments and laid the foundations of experimental psychology (Fröhlich 1929; Mollon & Perkins 1996).

27.2.1 Measurement of "sensation time"

In the 1920s and early 1930s several researchers, mainly in Germany, set out to measure the time it takes from a physical event to its perception. This time was referred to as *Wahrnehmungszeit* ("perception time," Hazelhoff & Wiersma 1924) or *Empfindungszeit* ("sensation time," Fröhlich 1923); today it would most likely be called the *perceptual latency* of an event.

To determine "sensation time" Fröhlich used a simple visual stimulus of an illuminated bar moving from behind an occluder. The basic finding was that the bar did not appear to start moving right at the edge of the occluder, but at a position a short distance farther ahead in its trajectory (Fröhlich 1923). Many variants of this basic finding were reported in a monograph (Fröhlich 1929). Fröhlich's assumption was that the time it takes the bar to move from the edge of the occluder to the first perceived position corresponds to the "sensation time" of the bar. He argued that it takes time for the bar to be processed, such that by the time the sensation is formed, the bar has moved on and therefore is visible in the shifted position. Early on this argument was criticized (Rubin 1929; Metzger 1932) for not making evident why after the "sensation time" has passed the bar should be seen in a forward shifted position.

Hazelhoff and Wiersma (1924), using a different methodology to determine the time of perception, had their observers track a moving fixation mark with ocular movements. A briefly flashed target was presented at the position of the moving mark when it had traversed half of the visual display. The position of this flash was misperceived as shifted in the direction of the eye movement. In this case, they argued that the position of the eye moves before the flash is processed and perceived, and the flash is perceived at the position the eyes point to after this "perception time" has passed.

Metzger (1932) accounted for the above findings as a special variant of the Hess phenomenon (1904). Stimuli of different contrasts have different "perception times" and are therefore perceived to move at different speeds. He explained the basic findings of Fröhlich with the assumption that, initially, a newly appearing object has a longer "perception time" than during later stages in its trajectory. In another experiment, Metzger tested the prediction that a bar already in motion would perceptually lead a bar starting from behind an occluder. This prediction was experimentally confirmed. Following from this Metzger studied an additional stimulus, which seems to be the earliest example of what would be considered a typical flash-lag display, although the term was coined much later by Nijhawan (1994). A single vertical slit is moved behind a cardboard occluder with two cutouts, one large horizontal cutout, where the moving slit is seen as a moving bar, and a small one above the large cutout, with the same width as the moving slit. Here the illuminated slit was seen as a flash (Fig. 27.2). Observers perceived the flash in the small cutout when the moving slit in the large cutout was already in a position further along the trajectory of movement. Metzger held that this phenomenon was in fact identical with Hazelhoff and Wiersma's earlier findings. Although Metzger's experiment did not involve any eye movements, the retinal stimulation was actually very similar to their earlier setup. In Hazelhoff and Wiersma's experiment, the moving object was stabilized on the retina by smooth pursuit eye movements, the flash was briefly presented next to the fixation position, and the background was moving continuously on the retina due to smooth pursuit eye movements.

Metzger concluded that both Hazelhoff and Wiersma (1924) and Fröhlich (1929) had measured the relative difference in the sensation time of a moving object and an abrupt onset. He deemed it impossible to measure absolute sensation time. His explanation of the flash-lag phenomenon is therefore in principle identical with later proposals of differential latencies for moving objects and flashes (Purushothaman et al. 1998; Whitney & Murakami 1998).

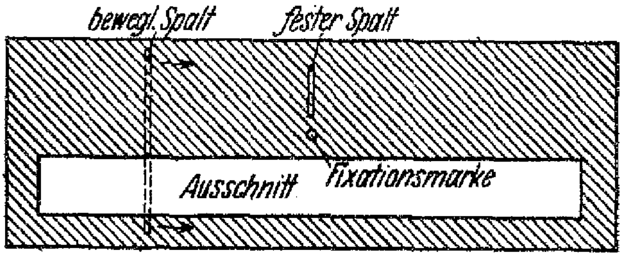


Fig. 27.2 To the best of our knowledge the first "flash-lag" stimulus display. A movable slit (*bewegl. Spalt*) is presented behind a cardboard occluder with two cutouts: a large horizontal cutout (*Ausschnitt*), where the slit is seen as a continuously moving object, and a small stationary slit (*fester Spalt*) above the fixation point (*Fixationsmarke*), where the moving slit appears as a brief flash. Metzger reported that observers saw the flash only when the continuously visible part of the slit was further ahead in the direction of motion. Source: Figure 3 from Metzger (1932).

27.2.2 Stroboscopically lit visual field

Unrelated to the reports of the flash-lag phenomenon from Germany in the early literature on perceptual latencies, another similar finding was published from England. MacKay (1958) observed that when the eyes are moved involuntarily, for example, by applying light pressure with the finger to the side of the eyeball, continuously lit objects in the visual field are perceived to move, whereas a stroboscopically lit background (at 5–6 Hz) does not. Instead it seems to "move sluggishly to 'catch up' with the self-luminous objects" (MacKay 1958). It was suggested that change and movement in the visual world have special significance and therefore are salient percepts, whereas the comparative amount of information in the stroboscopically lit field is not actually changing as much from one flash to the next.

Much later Mateeff and Hohnsbein (1988) replicated Metzger's first flash-lag finding and found differentially larger effects for motion toward the fovea compared to motion away from the fovea. These findings were also interpreted in terms of differential latencies.

27.2.3 The flash-lag effect in other features and modalities

The flash-lag effect has been measured using other visual features than motion, and cross-modally. Sheth et al. (2000) showed that other continuously changing features of an object are similarly "extrapolated" as position in the more standard motion flash-lag effect. When a visual stimulus patch continuously changes color, luminance, spatial frequency, or pattern entropy, and a second test patch with an identical feature value as in the changing patch is briefly flashed, observers perceive the continuously changing stimulus as further ahead in the dimension of change (Sheth et al. 2000). Although the degree of misalignment was highly variable depending on the feature in question, all features tested showed the forward shift.

Alais and Burr (2003) extended the flash-lag effect into the auditory domain. They found misalignments for both auditory frequency changes compared to a short burst of a single frequency, and auditory stereo position changes compared to a short tone from a static stereo position. Furthermore, when observers compared the position of an auditory stimulus moving in space to a visual flash, or a visual moving stimulus to an auditory burst, typical flash-lag results were obtained, with the moving stimulus cross-modally leading the "flashed" static stimulus (Alais & Burr 2003). These cross-modal flash-lag effects were found to be larger than unimodal effects within vision or audition. In other stimulus configurations, however, an additional auditory tone at the time of the flash can decrease the flash-lag effect, which might be the result of cross-modal temporal binding (Vroomen & de Gelder 2004).

Other cross-modal flash-lag effects have also been reported. When observers move their arm in the dark and compare the felt arm position to a visual flash ("motor flash-lag," Nijhawan & Kirschfeld 2003), or when observers compare the position of a moving tactile object on their forearm to a visual flash (Rojas-Anaya et al. 2005).

27.3 Flash-lag: theory and recent debates

27.3.1 The standard view and the ds-error postulate

Discussion of visual neural delays, the logically derived *ds*-error postulate (the assertion that there is a spatial lag in the perceived position of moving objects), and the flash-lag effect have appeared together time and again in recent articles (e.g., see Nijhawan 1994; Cavanagh 1997; Berry et al. 1999; Gegenfurtner 1999; Fu et al. 2001; Krekelberg & Lappe 2001). Fu et al. (2001) state: "Processing delay in the neural pathway on the order of tens of milliseconds should cause a significant offset between the perceived and the actual positions of a moving object. An appealing hypothesis is that the visual system can compensate for the neural delay and reduce the perceptual misalignment by extrapolating the trajectory of the moving object... Evidence for motion extrapolation originally came from the flash-lag illusion. . . ." In order to outline a coherent theoretical framework for the flash-lag effect, neural delays, and the broader topic of "neural representation of space and time," we present the following analysis and then go on to evaluate accounts of the flash-lag effect.

A human observer should perceive a flash after a significant delay relative to the actual time of the flash. It is unlikely that the nervous system can overcome this delay (van de Grind 2002). Likewise, one may reason that a moving object should be visible in a position it occupied in the recent past (*ds-error* postulate). Figure 27.3 presents this *standard view* graphically.

In 1994, Nijhawan suggested that the *ds*-error postulate was untenable. Other research on the flash-lag effect has also supported a revision of the *standard view*. The differential latency (Metzger 1932; Baldo & Klein 1995; also see Purushothaman et al. 1998; Whitney & Murakami 1998) and the spatial extrapolation (Nijhawan 1994) accounts of the flash-lag effect posit that the moving object is not perceived behind its actual position, but rather the percept of the moving object follows the physical object more closely than neural delays measured by discrete stimulation (e.g., by flashes) might suggest (Fig. 27.4).

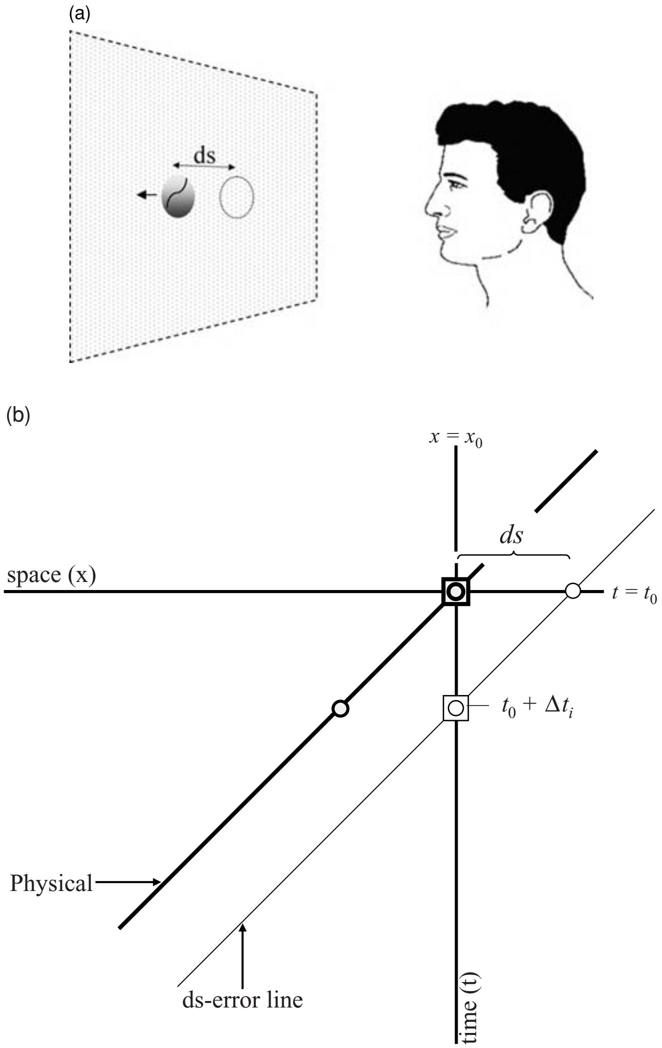


Fig. 27.3 (a) The standard view. The "average" observer views a ball moving leftward at constant velocity. At the instant depicted the physical position of the ball is shown by a distance *ds* ahead of the ball's perceived position (broken circle). This picture is a result of the *ds*-error postulate, which states that due to input delays in the visual pathways, the moving object should be visible where it was in the recent past. (b) The representation of the "standard" view with space–time plots of an object moving at constant velocity v (thick line) and how a neuron "sees" the object with some delay (thin line). A brief flash (square, thick outline) presented in position x_0 at time t_0 is seen (square, thin outline) by the same neuron in position x_0 at time $t_0 + \Delta t_i$, where Δt_i is input delay. Two events, the arrival of the moving object in position x_0 as seen by the neuron (thin outline circle on the $x = x_0$ line), and the physical arrival of the object in position x_0 (thick outline circle on $x = x_0$ line), occur at different times due to neural latency. At a given time (say t_0) there is spatial lag between the physical position of the moving object (thick outline circle on $t = t_0$ line) and the position in which the neuron "sees" the object (thin outline circle on $t = t_0$ line). The spatial lag between the physical position and the neurally represented position will be referred to as *ds-error*. The standard view asserts that the above is an accurate picture of neural delays and their impact on perception.

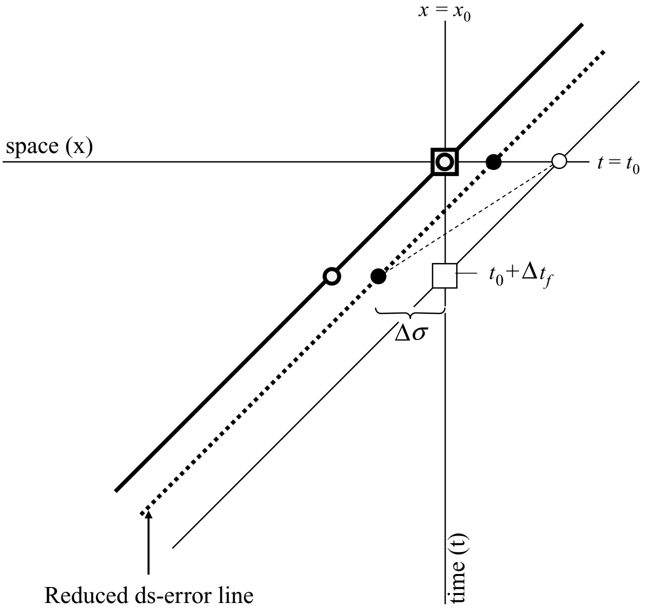


Fig. 27.4 A revision of the standard view is required by the flash-lag effect. The differential latency and spatial extrapolation accounts suggest that the percept of the moving object (filled black circle) is closer to the object's physical position than expected from measurements of neural delays with discrete stimuli-like flashes. So the thick broken ("*reduced ds-error*") line more accurately represents the perceived trajectory of the moving object. The *reduced ds-error* line is derived from the *ds-error* line by a shift corresponding to $\Delta\sigma$ = flash-lag effect. On the standard view a flash-lag effect could not occur unless the object appeared to speed up to cover the additional distance $\Delta\sigma$ (represented by thin broken line segment of different slope).

In the past decade and a half, the flash-lag effect has invited several accounts (Nijhawan 1994; Baldo & Klein 1995; Purushothaman et al. 1998; Whitney & Murakami 1998; Brenner & Smeets 2000; Eagleman & Sejnowski 2000; Krekelberg & Lappe 2000; Bachmann et al. 2003; Wojtach et al. 2008), and since the original publications (Nijhawan 1994, 1997; Baldo & Klein 1995; Khurana & Nijhawan 1995; Purushothaman et al. 1998; Whitney & Murakami 1998; Berry et al. 1999; Kirschfeld & Kammer 1999; Brenner & Smeets 2000; Eagleman & Sejnowski 2000; Krekelberg & Lappe 2000; Schlag et al. 2000; Sheth et al. 2000) these accounts have undergone various revisions and generalizations. Here for purposes of analysis we partition the accounts under three headings: Visual Prediction; Spatiotemporal Integration, Sampling, & Postdiction; Differential Latency & Attention.

Based on various assumptions and theoretical constructs used, the accounts may be separated into three different categories: (1) accounts that place the perceived object on the **thick** broken (*reduced ds-error*) line, throughout some extended motion trajectory (namely, differential latency and visual prediction), (2) accounts that place the perceived object on the **thick** broken line for a short period of the motion trajectory (namely, postdiction account,

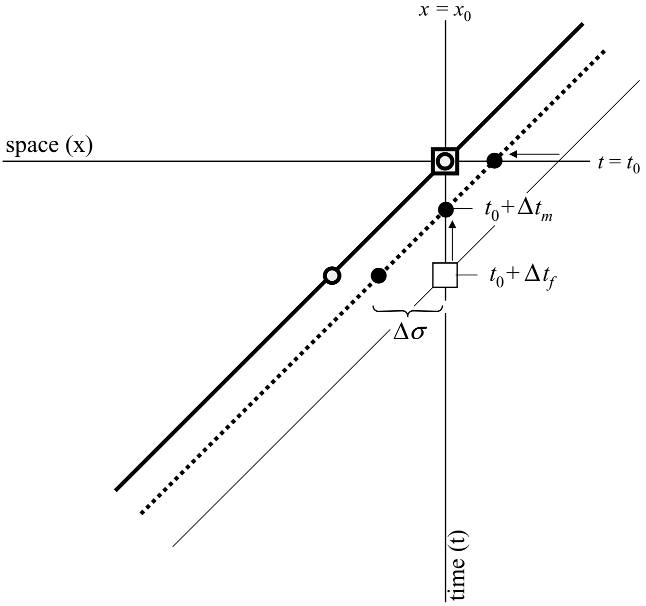


Fig. 27.5 The differential latency account places the moving object on the thick broken line as it assumes that moving objects are processed with a shorter latency (Δt_m) in relation to the latency (Δt_f) with which flashes are processed. This reduced latency, depicted by the upward pointing vertical arrow leading to the filled black circle, places the moving object on the *reduced ds-error* line (the thick broken line). Latency reduction occurs for any extended motion trajectory. As this account is strictly temporal the shift in coordinate must be parallel to the vertical (time) axis. The spatial extrapolation account assumes a spatial displacement that places the perceived object on the *reduced ds-error* line, depicted by the leftward pointing horizontal arrow leading to the filled black circle. Like differential latency spatial extrapolation occurs for any extended motion trajectory. Adapted from Krekelberg and Lappe (2001).

see the following), and (3) accounts that place the perceived object on the **thin** continuous (*ds-error*) line throughout (namely, position-averaging and motion-sampling accounts; see the following). Although the visual prediction and differential latency accounts both place the moving object on the *reduced ds-error* line throughout, the reasons they do so are different (see Fig. 27.5 and text following).

27.3.2 Visual prediction

Visual prediction holds that motion is of fundamental importance to the animal. The visual system is designed to exploit the occurrence of motion at constant speeds in fixed directions over short periods of time to predict a moving object's position. This prediction is not of future position, but of the present position (Cavanagh 1997), which achieves the compensation of visual processing delays. The moving object is therefore not seen in a position lagging behind the physical position, in contrast to the *standard view*, but closer to its real position at any point along its trajectory. Because the flash is a sudden unpredictable event, neural delays in its registration cannot be compensated for in the same way. In the time it takes the visual system to process the flash, the moving object has moved on, which gives rise to the flash-lag effect (Nijhawan 1994).

27.3.2.1 Laying-to-rest the ghosts of flash-lag past and flash-lag future

For a number of years, the results of the following two experimental conditions seriously undermined the viability of the visual prediction account of the flash-lag effect: (1) The flash-terminated condition (originally called the "past-interval," see Nijhawan 1992 in appendix) consists of events only up to and inclusive of the flash. In this condition the flash is aligned with a moving object's last seen position. In other words, the moving object disappears from view simultaneous with flash offset. (2) In the flash-initiated condition (originally called the "future-interval," see Nijhawan 1992 in appendix), the flash is aligned with the first position of a moving object. The moving object appears simultaneous with flash onset. The counterintuitive findings are as follows: the flash-terminated condition produces no flash-lag effect, whereas the flash-initiated condition produces an effect equal in magnitude to that observed with the standard full motion display (see Nijhawan 1992, in Appendix; Khurana & Nijhawan 1995).

Although intuitively these results are inconsistent with, and would even seem to rule out the visual prediction account, intuitions can be overturned in the presence of new experimental investigations. Recent analyses and experimental findings related to the flash-terminated and flash-initiated conditions (Maus & Nijhawan 2006, 2008, in press; Nijhawan 2008a) have not only made visual prediction compatible with these results but actually even more viable (also see Nijhawan 2002). Briefly, the transient signals produced by the abrupt disappearance of the moving object in the flash-terminated condition eliminate the flash-lag effect. These signals, which are present for any significant velocity changes (not just disappearances), suppress the perception of the forward shift. Indeed elimination of transients brings back the forward shift (Maus & Nijhawan 2006, 2008; also see Section 27.3.3.1 "Empirical challenges for the integration and sampling accounts" following).

In the flash-initiated condition, the intuition that extrapolation should take a significant time to be initiated after motion onset is wrong by about two orders of magnitude! Even by conservative estimates, the time taken for an extrapolated percept to be produced can be less than 2% of the 100 msec baseline delay typically thought to exist between stimulation of the retina and perception (De Valois & De Valois 1991; Nijhawan 2008a). Given that the required time for spatial extrapolation is so small, motion extrapolation is an extremely efficient mechanism that could start and finish virtually anytime during the required baseline (100 msec) delay after motion onset (Khurana & Nijhawan 1995). For a thorough analysis of time requirements of spatial extrapolation see Nijhawan (2008a, Section 5.2.1, page 185).

27.3.2.2 Neurophysiological support for visual prediction

The earliest stage of visual processing is in the retina itself. Interestingly, prediction can already be found in retinal ganglion cells, which form the output layer of the retina. In response to a moving stimulus, the peak-firing rate in the ganglion cell population is found to be at or even ahead of the leading edge of the stimulus (Berry et al. 1999). There is no anatomical input from higher areas in the visual pathway back to the retina, so this predictive forward shift of the neural representation must come about due to local interactions in the retina.

Neural activity in lateral geniculate nucleus (LGN) is highly shaped by cortical feedback. In fact, although only ∼10% of synaptic input into LGN cells stems from the retina, far more – about 30% – stems from cells in layer 6 of V1, which in turn receive fast feedback from motion area MT/V5 (Sillito et al. 2006). Anatomically, the feedback from V1 to the LGN, though local, spreads to nearby retinotopic locations along the orientation of the receptive field and in the direction of motion selectivity of the V1 cell. The spread of feedback to other retinotopic locations in the LGN could cause a shift in the retinotopic position of activity toward future stimulus positions. In experiments manipulating the responsiveness of MT/V5 by local drug application in macaque cortex, significant changes in the response properties of cells in V1 and the LGN have been shown (Sillito & Jones 2002). This feedback loop from MT/V5 via V1 to the LGN has been interpreted to "enable predictive modulation of circuitry at earlier levels, *even before* (italics added) the input arrives" (Sillito et al. 2006, p. 307).

Directionally selective neurons exist in primate V1 and other visual cortical areas. Studies have shown position shifts in the neural activity in retinotopic maps that could underlie predictive mislocalization phenomena. Sundberg et al. (2006) convincingly showed a distortion of retinotopy in macaque area V4 for stimuli that involved a color singleton in an apparent motion sequence (Cai & Schlag 2001). Neurons responded to an odd-colored "flash" even before it entered the neuron's receptive field. This shift of the receptive field also occurred when the odd-colored flash was the final presentation of the bar in apparent motion, in contrast to human observers not showing a perceptual shift of the perceived position (Sundberg et al. 2006). This finding is consistent with the idea of neurons predictively coding positions in the visual field.

27.3.3 Spatiotemporal integration, sampling, and postdiction

Motion averaging (Krekelberg & Lappe 2000), motion sampling (Brenner & Smeets 2000), and postdiction (Eagleman & Sejnowski 2000) accounts all begin with the premise that the percept of a moving object lags behind the object's physical instantaneous position. In other words, these accounts do not subscribe to the above-proposed revision of the *standard view* (see Fig. 27.4). On the spatial-averaging account, the visual system averages the position of the moving object over some temporal window. Thus the precise position of the moving object becomes available some time after the perception of the flash (Fig. 27.6). The motion-sampling account offers a similar explanation. Another variant of temporal integration has been advanced in which averaging is weighted toward the most recently sampled positions (Roulston et al. 2006). This is offered as an explanation for a

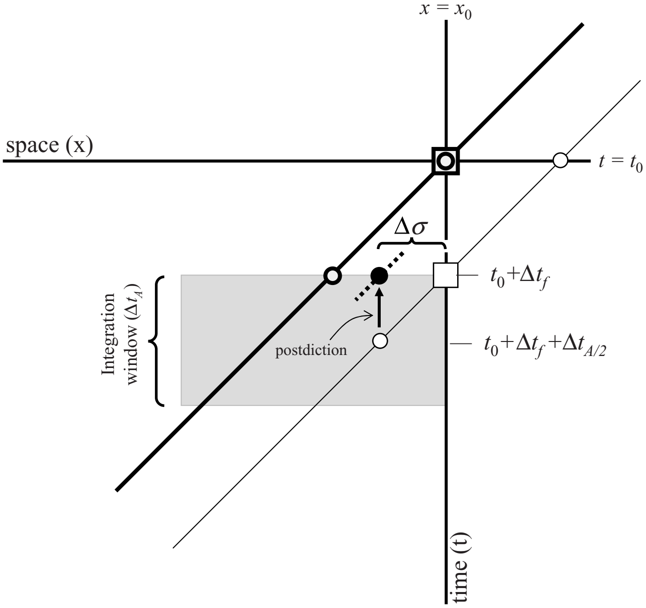


Fig. 27.6 Accounts that do not subscribe to a revision of the standard view claim that the thick broken (*reduced ds-error*) line in Figure 27.4 does *not* describe the correct trajectory of the perceived moving object. These accounts either place the object on the thin continuous (*ds-error*) line throughout (spatial-averaging and motion-sampling accounts) or on the thick broken line briefly (postdiction account). On the spatial-averaging account the moving object's perceived position is based on an average of the moving object's position in a certain time window (filled light gray rectangle of height Δt_A). Thus the precise position of the moving object (thin outline circle) becomes available after some time ($\Delta t_{A/2}$) of the perception of the flash. The "integration and postdiction" account assumes further that the output of integration (similar to averaging) is "postdicted" to the time the flash is registered (vertical arrow leading to filled black circle shown on short segment of thick broken line).

small undershoot of the final perceived position of a moving object (Baldo et al. 2002; Roulston et al. 2006), as well as the Fröhlich and standard flash-lag effects.

Both motion averaging and motion sampling accounts assume that the temporal mismatch (= ½ the temporal window) between the perception of the flash and output of the averaging processes is not so relevant, as small temporal asynchronies are not detected by the visual system. The postdiction account (Eagleman & Sejnowski 2000) suggests that the output of the motion integration process (similar to averaging) is "postdicted to the time of the flash" (p. 2038), depicted by the upward pointing vertical arrow leading to the filled black circle (Fig. 27.6).

Motion averaging and integration in the visual system is not only well supported by various psychophysical observations, but also by commonly seen behavior of neurons. In response to change in stimulation, neurons do not increase or decrease their activity level instantaneously. For example, retinal ganglion cells continue to be active for a significant duration after cessation of stimulation (Kratz & May 1990). If persisting neural output were "directly" used by the visual system in determining a moving object's instantaneous position, then this representation would be a "fuzzy" indicator of the object's physical position. So, position averaging is necessary for accurate localization. Thus position averaging is not just a likely neural mechanism for more precise visual localization of moving objects (Burr & Morgan 1997) but also necessary for visually guided behavior, such as catching moving objects for which accurate localization is essential. Given such a strong basis for position averaging, it should take place whenever the animal encounters motion in the world.

So the first challenge for the position averaging account is: Why does the averaging start after $t_0 + \Delta t_f$, that is, after the registration of the flash (Fig. 27.6)? In other words, why is the lower bound of the integration window aligned with $t_0 + \Delta t_f$ in Fig. 27.6 (and in Figure 2D of Krekelberg & Lappe 2001)? According to the logic outlined above, a moving integration window, whose output is also moving, better captures the integration process. If position averaging took place for moving objects throughout their trajectory, then the output of averaging should be available not just at $t_0 + \Delta t_f + \Delta t_{A/2}$ (as required to explain flash-lag) but also at the earlier time when the flash is first registered at $t_0 + \Delta t_f$. (This output would be based on motion input, which was presumably there prior to the flash). A moving output of the integration process, carried out over a fixed window height, should be aligned with the flash. Thus, the perceived position (given by the average) of the moving object at the time the flash is first perceived should be aligned with the flash, thus rendering a lack of a flash-lag effect. In that first instant, there shouldn't be a lag, but rather it should "develop" in the time immediately following $t_0 + \Delta t_f$ (during the time the flash visibly persists). This is contrary not just to the phenomenological observations of the flash-lag effect (Nijhawan 1994), where observers claim that the flash lags the moving object when first seen, but would render certain flash-lag results impossible. For example, the color decomposition result (Nijhawan 1997), in which a red flashed line superimposed on a moving green bar appears to lag behind the bar and appears red (as opposed to yellow), critically depends on the flashed line appearing in a lagging position when first seen. Indeed, inconsistencies are seen in Krekelberg and Lappe's (2001) Figure 2D (and its adaptation, our Fig. 27.6). These figures show a continuous delayed line, which by definition depicts the moving object's registered positions between times $t_0 + \Delta t_f$ and $t_0 + \Delta t_f + \Delta t_{A/2}$ (and earlier), and yet the account works only if the *instant* corresponding to $t_0 + \Delta t_f + \Delta t_{A/2}$ is selected. The figure, in fact, shows that there should be no effect at $t_0 + \Delta t_f$, which counters the very definition of flash-lag.

Further challenges arise for the postdiction account due to the notions of "flash reset" and "postdiction to the time of the flash." On the postdiction view (Eagleman & Sejnowski 2000), the flash "resets" motion integration (p. 2037), and the output of motion integration (similar to averaging) is "postdicted to the time of the flash" (p. 2038). Clearly, for this account flashes are critical. However, there are numerous forward-shift motion phenomena that do not involve flashes, so for such phenomena the concepts of "flash-reset" and "postdiction to the time of the flash" do not logically arise.

More problematic, however, is the notion that the visual system "postdicts" the output of motion integration to the time of the flash registration. In other words, the "postdicted" object (based on motion integration) should lie on the horizontal line passing through $t_0 + \Delta t_f$. But this point is on the *reduced ds-error* line (see thick broken line segment in Fig. 27.6). The question then is what happens to this postdicted object when motion continues after the flash. Does the object now continue its trajectory as given by the thick broken (*reduced ds-error*) line, or does the object trajectory return to the delayed thin continuous (*ds-error*) line? (The postdiction account assumes that prior to the flash, and sometime after it, the *ds-error* line accurately reflects the perceived positions of the moving object.) According to the postdiction account, the object trajectory cannot continue to be the thick broken line as then, after the flash (for example, if a second flash is presented in alignment with the moving object), this account will become indistinguishable from the differential latency and spatial extrapolation accounts. On the other hand, if the object trajectory returned to the thin continuous (*ds-error*) line, then the object would need to slow down (converse of the process depicted by the thin broken line segment in Fig. 27.4). Because this predicted slowdown would be easy to note, and has never been noted, this account clearly cannot be an explanation of the flash-lag effect.

27.3.3.1 Empirical challenges for the integration and sampling accounts

The postdiction account claims that future events (after the flash) impact the position of the moving object. Likewise, the sampling account claims the flash-lag effect is based on the speed of motion after the flash (Brenner & Smeets 2000). In addition to the above-mentioned logical challenges faced by these accounts, two recent studies provide evidence against the claim that the forward shift of a moving object is influenced by information from the future trajectory of the moving object. The first study consisted of two experimental conditions (Maus & Nijhawan 2006). In the first condition, a stationary white dot was presented in a series of positions P_1, P_2, $P_3, \ldots,$ $P_{N\text{-}1}$, $\mathrm{P_N}$ (on an imaginary circle), such that the dot intensity at $P_1 >$ dot intensity at $P_2 >$ dot intensity at $P_3 >$ dot intensity at P_4, and so on. At P_N the dot was invisible. Between positions P_1 and P_N a position was determined, for each observer, where the dot's intensity was at the observer's absolute detection threshold (defined as 50% detectability). This position was called the "edge of visibility," P_{Tstat} (detection threshold position for the stationary dot). Note: In positions just beyond (after) P_{Tstat} the dot did not provide sufficient input to lead to the observer seeing the dot reliably.

In the second condition, a smoothly moving dot underwent exactly the same intensity decrement as the dot in the first condition; the intensity-position relationship was identical to the first condition. Is the "edge of visibility" in the motion condition (P_{Tmov}) the same as P_{Tstat}, or is it shifted forward? In other words, does the smoothly moving dot perceptually disappear for the observer at P_{Tstat}, or is the disappearance position shifted forward? We found that P_{Tmov} was shifted by 175 msec relative to P_{Tstat}, in the direction of motion. Thus the dot in the motion condition was visible in positions where the stationary dot could not be seen because the latter's intensity was below detection threshold in those positions. This shift cannot be caused by input from the dot in positions beyond P_{Tstat}, as this position was (by definition) the point beyond which input was insufficient to yield visibility of the dot. So, this forward shift can only have occurred due to the input from the past positions of the moving dot before it reached P_{Tstat}.

In a second study, a smoothly moving line segment moved from the signaling part of the retina toward the retinal blind spot and then crossed into the blind spot (Maus & Nijhawan 2008). Where do observers perceive the line to disappear in the visual field? In this case, again, it was found that the line segment was visible in positions past the boundary of the blind spot in regions well into the blind spot. Once again, in this case the forward shift was seen, but the stimulus input from the future (later) positions was missing. Clearly, there is no input from the blind spot. So any visibility of the moving line in the blind spot region could only be based on earlier input from the moving object, before it reached the blind spot boundary.

It is clear from both these studies that the forward shift in the position of moving objects, at least in the stimuli studied, could not be due to later motion signals being integrated with earlier motion signals. Earlier motion signals impacting the position of the moving object parsimoniously explains not just the results of the above-mentioned studies but also flash-lag results, including flash-initiated and flash-terminated results (as discussed in the section on visual prediction, see previous).

27.3.4 Differential latency and attention

27.3.4.1 Movement of attention

Soon after the original announcement of the flash-lag effect (Nijhawan 1994), an alternative to motion extrapolation was proposed by Baldo and Klein (1995). They showed that the illusory forward displacement of rotating moving dots relative to the perceived positions of flashed dots depends on the eccentricity of the flashes, with larger eccentricity leading to larger forward displacements. It was proposed that the time for the flashes to be perceived depends on their eccentricity, as it takes time to reallocate attentional resources to the more peripheral positions of the flashes. The flash-lag effect is thus a consequence of the time cost in shifting attention from the moving object to the flash and back. "Cue-induced visual focal attention" has been suggested as a mechanism to accelerate perceptual processing of a moving object along its trajectory (Kirschfeld & Kammer 1999), and thus to cause the flash-lag effect.

In reply to Baldo and Klein (1995), Khurana and Nijhawan (1995) created a similar stimulus in which flashing and moving objects were not presented in distinct spatial locations. In a rotating line of dashes, additional dots were flashed in positions interjected between the moving dashes, and the flash-lag effect was still perceived. Furthermore, when the flash was presented at the time the moving object was set into motion (the flash-initiated condition), an unabated flash-lag effect was perceived (Khurana & Nijhawan 1995). In this flash-initiated condition, attention is equally oriented to both of the two objects (moving or flashed), so there is no attentional disadvantage for one or the other. These findings argue against an explanation of the flash-lag effect solely in terms of the delays associated with attention shifts.

In the above studies, (Baldo & Klein 1995; Khurana & Nijhawan 1995) allocation of attention was inferred. A later study explicitly manipulated and measured attention in the flash-lag paradigm by directing attention either toward or away from the positions of flashes (Khurana et al. 2000). Although simple reaction times to flashes at cued locations were quicker than for uncued locations, there was no difference in the magnitude of the flash-lag effect as a function of attentional modulation. These findings further argue against an explanation of the flash-lag effect based on delays in the redistribution of attentional resources from moving to flashed stimuli. In other studies, it has been shown that the effect size can be modulated by voluntary attention (Baldo et al. 2002; Namba & Baldo 2004). It is likely that attention here influenced perceptual latencies, but although attention is not the underlying cause of the flash-lag effect, it may be an additional influence (Namba & Baldo 2004).

27.3.4.2 Differential latency

Perhaps the simplest account of the flash-lag effect is the differential latency account (Metzger 1932; Purushothaman et al. 1998; Whitney & Murakami 1998). This view suggests that the percept of the moving object lags the object's physical position by a distance $v\Delta t_m$, which is less than $v\Delta t_f$; here Δt_m and Δt_f are the latencies for the moving and the flashed object, respectively. The flash-lag effect reflects the relationship $v(\Delta t_f - \Delta t_m) = \Delta\sigma$, where $\Delta\sigma$ is the flash-lag effect. This account has great intuitive appeal, as it is well known that different stimuli are processed with different delays.

The main weakness of the differential latency account is that $\Delta t_f - \Delta t_m$ must equal about 80 msec, the typically cited value of $\Delta\sigma/v$ (e.g., see Eagleman & Sejnowski 2000). This is a large latency difference that is a priori unlikely as flashes are extremely effective at stimulating the visual system. Furthermore, such a large difference should have been revealed and well established owing to the innumerable neurophysiological studies over the last century that have employed both static flashes and moving objects to stimulate the visual system. Failures to find differences usually go unreported; however, some articles have suggested that, if anything, neural response to flashes occurs slightly faster than to motion (Raiguel et al. 1989; however, see Jancke et al. 2004). Furthermore, the "differential latency" account is ruled out as a complete explanation of the flash-lag effect by the fact that the flash-initiated condition (Khurana & Nijhawan 1995), in which the moving object comes on simultaneously with the flash, produces an undiminished flash-lag effect. How can a moving object that suddenly appears on the screen be perceived more quickly than a flash, and that too by 80 msec? If this were the case, the automobile industry would have come up with a brake design, which would turn on a moving stimulus instead of turning on a stationary light! We note, again, that the spatial extrapolation account does not face this speed-of-processing challenge posed by the flash-initiated condition, as a spatial shift in coordinates of neural activity representing the moving object incurs a trivial added delay relative to baseline delay (see Nijhawan 2008a, Section 5.2.1, p. 185).

Not surprisingly several psychophysical studies have failed to find any reliable latency difference between moving objects and flashes (Eagleman & Sejnowski 2000; Alais & Burr 2003; Nijhawan et al. 2004; Ichikawa & Masakura 2006). At this juncture, the only safe conclusion can be that if there is a latency advantage to the processing of moving objects, in relation to flashes, then this advantage is probably small in comparison to $\Delta\sigma/v$ (=80 msec).

27.3.5 Comparison of spatial and temporal accounts

There is no doubt that the variables of differential latency and attention do play a role in the flash-lag effect. But their contributions are unlikely to be sufficient to explain the entire effect. A spatial account such as motion extrapolation fares better. An additional advantage of the motion extrapolation account is that it potentially explains not just the flash-lag effect but also numerous other forward-shift phenomena related to motion per se. There is also neurophysiological support for this account (discussed previously). But perhaps the greatest appeal of the motion extrapolation account is that it presents some remarkably simple modifications of existing facts and ideas to yield new insights into the connections between neurophysiology and perception. We consider two such insights here. One concerns a modification of the notion of a motion sensor to incorporate spatial extrapolation while adhering to the notion of retinotopic representation of visual space. The second introduces the notion of a "diagonal neural pathway."

Among the modeling achievements concerning visual neural processes, the concept of a directional motion sensor stands out for its simplicity and elegance. A motion sensor, connected to two input lines with a delay in one of the lines, produces a response to motion in one direction (rightward in Fig. 27.7(a)) but not to motion in the opposite direction. Of all neural–perceptual isomorphisms, the concept of topographic retinal maps and their contribution to perception of visual space has been most fundamental to neurophysiology since retinotopic maps were first discovered. A neurally simple modification of the concept of the directional motion sensor is suggested by motion extrapolation that allows for compensation of neural delays while keeping the spatial isomorphism unaffected. Figure 27.7(b) shows a modified motion sensor that is not just a motion sensor but also provides location information of the moving object. This motion sensor is shifted in the direction for which it is tuned. The shift compensates for the transmission delay.

One interesting further possibility is revealed by this modified scheme. Two oppositely tuned directional motion/position sensors together could serve to localize stationary objects (Fig. 27.7(c)). This suggests that motion may be the medium that sets up retinotopic maps initially during development, for which there is evidence (Meister et al. 1991). These very maps are later used by the mature system for localizing both moving and stationary objects (Nijhawan 2008b).

The notions of horizontal and vertical pathways are well established in neurophysiology. Motion extrapolation leads naturally to considering a "diagonal pathway." Consider a simple two-layered network in which an input layer of neurons is connected to an output

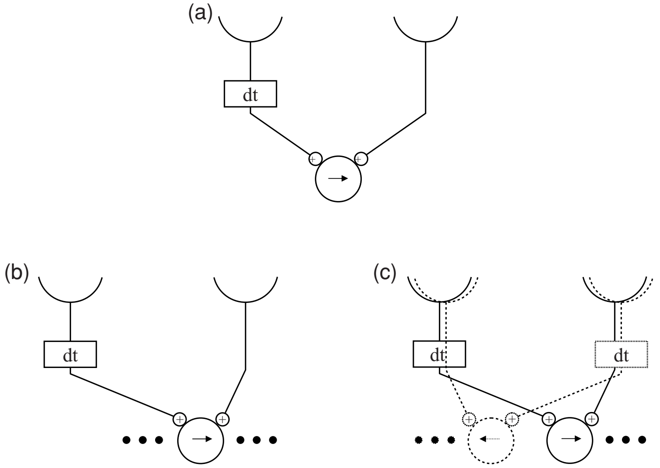


Fig. 27.7 (a) Shows the standard Reichardt-type directional motion sensor. (b) A modified sensor that also senses object position. When this sensor is active it gives a forward-shifted position (relative to that in (a)) to compensate for neural delays. (c) A pair of modified motion sensors that are oppositely tuned for motion direction. The output of this pair would signal position of stationary objects, which are caused to shift over the retina by small (oscillatory) eye movements.

(perceptual) layer. The input layer has two opposing requirements. One is to transmit information about significant world events onto the output layer as quickly as possible, whereas the second is to process the input so that it is usable by later neurons, so the animal is able to produce appropriate response. This processing is what slows down transmission (Nijhawan & Wu 2009). (As a contrast one could imagine transmission of raw data to spring-loaded effectors that reacted as quickly as possible.) In the case of moving objects, the visual system solves the problem of neural delays by sending information along diagonal pathways, which are made up of horizontal and vertical neural segments (e.g., in the retina the horizontal cells constitutes the horizontal segment, and the bipolar cells constitute the vertical segment; the two together make up the "diagonal" pathway). Consider a two-layered network with sensory neurons X_1, X_2, ... connected to perceptual neurons X'_1, X'_2, ... Neuron X_1 is connected to neuron X'_1, neuron X_2 is connected to neuron X'_2, and so on. S_1, S_2, and so on specify successive locations of a moving stimulus.

$$\begin{array}{ccc} S_1 & X_1 & X'_1 \\ S_2 & X_2 & X'_2 \\ S_3 & X_3 & X'_3 \\ . & . & . \\ . & . & . \\ . & . & . \end{array}$$

In order to compensate for the transmission delay, such a network has to be tuned to the delay such that when the stimulus arrives at S_2 (stimulating X_2), X_2 sends input "diagonally" to X'_3 (not to X'_2) so that information gets to X'_3 simultaneously with the stimulus getting to S_3 (Nijhawan & Wu 2009).

27.4 Conclusion

Over the past decade the notion of visual prediction has led to a lively debate. The appeal of visual prediction is that on this view visual mechanisms are assigned the task of estimating *and* compensating for visual delays. Visual prediction opposes the *standard view* (the *ds-error* postulate; the notion that neural delays should cause a lag in the moving object's perceptual position relative to its physical position). New analyses (Nijhawan 2008a) and data (Maus & Nijhawan 2006, 2008) suggest that empirical results with unpredictable motion, in particular the flash-terminated and flash-initiated conditions, are compatible with visual prediction (also see Nijhawan 2002). In fact, the times involved by the putative extrapolation mechanism are so short that *ds-error* could be reduced to zero (or even have a negative value due to "overextrapolation," see Nijhawan 2008a).

On a final philosophical note, both the notions of differential latency and spatial extrapolation attempt to reduce the discrepancy between the "real" world and the "experienced" world, but there is a deep conceptual difference between the two accounts. Differential latency starts by assuming a "real" world and a delayed representation of it by the brain. Most of the thinking in psychophysics and neuroscience represents this view. Differential latency proposes that when possible the "reconstructed world" follows the "real" world closely. Visual prediction starts with a fundamentally different assumption. It assumes that neural representations (and not the "real" world) exist first. The notion of the "real" world emerges from these representations. This proposal is concordant with that of Ernst Mach (see Ratliff 1965, for a discussion of Mach's position). The use of the term "real" world, on this view, is justified only for everyday practical affairs. A most illuminating statement reflecting this challenge is by the Gestalt psychologist Wolfgang Köhler (1947): "If a wound is not the gun which emitted the projectile, then the thing which I have before me, which I see and feel, cannot be identical with the corresponding physical object."

In Mach's terminology, the complex of neural representations leads to the concept of the "real" world. By reversing this argument Mach produced a most informed critique of Kant's "thing in itself." Mach wrote: "The vague image that we have of a given permanent complex, being an image that does not perceptibly change when one or another of the component parts is taken away, gradually establishes itself as something that exists by itself. Inasmuch as it is possible to take away singly every constituent part without destroying the capacity of the image to stand for the totality and of being recognized again, it is imagined that it is possible to subtract all parts and to have something still remaining. Thus arises the monstrous notion of a thing in itself, unknowable and different from its 'phenomenal' existence" (Mach 1885/1897, pp. 5–6). Likewise, the impetus for the motion extrapolation view comes from a desire to remove the "vague image," a ghost that is unknowable (because all neural processes leading to knowledge are subject to delays) but that, nonetheless, trails behind the phenomenal object in motion.

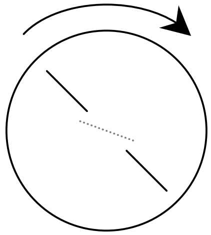
Appendix

1415 – 7:15

MISALIGNMENT OF CONTOURS THROUGH THE INTERACTION OF APPARENT AND REAL MOTION SYSTEMS.

Romi Nijhawan, Psychology Department, University of California Berkeley, CA

Observers viewed a rotating diameter of a physical disc whose central third segment flashed (for 5 msec) at predetermined times while the two, outer third segments, were illuminated continuously. Thus, the latter segments underwent real motion while the former flashed at rates which did or did not produce apparent motion. The outcome was quite unusual. The flashed segment appeared rotated by up to 10 degrees (for angular velocity of diameter = 38 rev/min) in the direction opposite to the rotation of the continuous segment (see Fig). The effect was directly related to the angular velocity of the diameter. When the flashing rate was adjusted so that the central segment underwent apparent rotation, the effect was reduced and was inversely related to the rate of flashing. Three possible explanations were considered: 1) Delayed processing of the flashed stimulus, 2) movement-induced positional bias, and 3) visible persistence in the flashed segment and the lack thereof in the continuously moving segments. To decide between these possibilities the display (for a flash rate = 1 flash/rev) was separated into two temporal intervals. One interval consisted only of events before the flash (past-interval) and the other interval only of events after the flash (future-interval). The observers reported no misalignment effect in the past-interval but did so in the future-interval. The magnitude of the effect for the future-interval was as strong as that for the past+future interval. This result supports the third account above indicating that motion perception is necessary to erase motion smear and that motion is computed prior to form.



– *Invest. Ophthal. Vis. Sci.* ***33**, 1415, 1992*

References

Alais, D., & Burr, D. (2003). The "flash-lag" effect occurs in audition and cross-modally. *Current Biology* **13**(1): 59–63.

Bachmann, T., Luiga, I., Põder, E., & Kalev, K. (2003). Perceptual acceleration of objects in stream: evidence from flash-lag displays. *Conscious Cogn* **12**: 279–297.

Baldo, M. V. C., Kihara, A. H., Namba, J., & Klein, S. A. (2002). Evidence for an attentional component of the perceptual misalignment between moving and flashing stimuli. *Perception* **31**: 17–30.

Baldo, M. V., & Klein, S. A. (1995). Extrapolation or attention shift? *Nature* **378**: 565–566.

Berry, M. J., Brivanlou, I. H., Jordan, T. A., & Meister, M. (1999). Anticipation of moving stimuli by the retina. *Nature* **398**(6725): 334–338.

Brenner, E., & Smeets, J. B. J. (2000). Motion extrapolation is not responsible for the flash-lag effect. *Vision Res* **40**: 1645–1648.

Burr, D. C., & Morgan, M. J. (1997). Motion deblurring in human vision. *Proc R Soc Lond B Biol Sci* **264**(1380): 431–436.

Cai, R., & Schlag, J. (2001). A new form of illusory conjunction between color and shape [Abstract]. *J Vis* **1**(3): 127. Retrieved from doi:10.1167/1.3.127

Cavanagh, P. (1997). Predicting the present. *Nature* **386**(6620): 19, 21.

De Valois, R. L., & De Valois, K. K. (1991). Vernier acuity with stationary moving gabors. *Vision Res* **31**: 1619–1626.

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science* **287**(5460): 2036–2038.

Fröhlich, F. W. (1923). Über die Messung der Empfindungszeit. *Zeitschrift für Sinnesphysiologie* **54**: 58–78.

Fröhlich, F. W. (1929). *Die Empfindungszeit*. Jena: Gustav Fischer.

Fu, Y. X., Shen, S. Y., & Dan, Y. (2001). Motion-induced perceptual extrapolation of blurred visual targets. *J Neurosci* **21**(20): (RC172).

Gegenfurtner, K. (1999). Neurobiology. The eyes have it! *Nature* **398**(6725): 291–292.

Ghez, C., & Krakauer, J. (2000). The organization of movement. In E. R. Kandel, J. H. Schwartz, & T. M. Jessel (eds.), *Principles of Neural Science* (4th ed.). New York: McGraw-Hill.

Hazelhoff, F., & Wiersma, H. (1924). Die Wahrnehmungszeit. *Zeitschrift für Psychologie* **96**: 171–188.

Hess, C. (1904). Untersuchungen über den Erregungsvorgang im Sehorgan bei kurz- und länger dauernder Reizung. *Pflügers Archiv für die gesamte Physiologie* **101**: 226–262.

Ichikawa, M., & Masakura, Y. (2006). Manual control of the visual stimulus reduces the flash-lag effect. *Vision Res* **46**(14): 2192–2203.

James, W. (1890/1952). *The Principles of Psychology*. London: Encyclopaedia Brittanica.

Jancke, D., Erlhagen, W., Schoner, G., & Dinse, H. (2004). Shorter latencies for motion trajectories than for flashes in population responses of cat primary visual cortex. *J Physiol* **556**(Pt 3): 971–982.

Jordan, M. I. (1995). Computational motor control. In M. S. Gazzaniga (ed.), *The Cognitive Neurosciences*. Cambridge, MA: MIT Press.

Khurana, B., & Nijhawan, R. (1995). Extrapolation or attention shift: reply. *Nature* **378**(6557): 565–566.

Khurana, B., Watanabe, K., & Nijhawan, R. (2000). The role of attention in motion extrapolation: are moving objects “corrected” or flashed objects attentionally delayed? *Perception* **29**(6): 675–692.

Kirschfeld, K., & Kammer, T. (1999). The Fröhlich effect: a consequence of the interaction of visual focal attention and metacontrast. *Vision Res* **39**(22): 3702–3709.

Köhler, W. (1947/1992). *Gestalt Psychology: An Introduction to New Concepts in Modern Psychology*. New York: Liveright.

Kratz, K. E., & May, J. G. (1990). Response persistence of cat retinal ganglion cells to the temporally discrete presentation of sinewave gratings. *Int J Neurosci* **52**(1–2): 111–119.

Krekelberg, B., & Lappe, M. (2000). A model of the perceived relative positions of moving objects based upon a slow averaging process. *Vision Res* **40**(2): 201–215.

Krekelberg, B., & Lappe, M. (2001). Neuronal latencies and the position of moving objects. *Trends Neurosci* **24**(6): 335–339.

Mach, E. (1885/1897). *Contributions to the Analysis of the Sensations* (C. M. Williams, Trans.). Chicago: The Open Court Publishing Company.

MacKay, D. M. (1958). Perceptual stability of a stroboscopically lit visual field containing self-luminous objects. *Nature* **181**: 507–508.

Mateeff, S., & Hohnsbein, J. (1988). Perceptual latencies are shorter for motion towards the fovea than for motion away. *Vision Res* **28**(6): 711–719.

Maus, G. W., & Nijhawan, R. (2006). Forward displacements of fading objects in motion: the role of transient signals in perceiving position. *Vision Res* **46**(26): 4375–4381.

Maus, G. W., & Nijhawan, R. (2008). Motion extrapolation into the blind spot. *Psychol Sci* **19**(11): 1087–1091.

Maus, G. W., & Nijhawan, R. (in press). Going, going, gone: localizing abrupt offsets of moving objects. *J Exp Psychol Hum Percept Perform.*

Meister, M., Wong, R. O., Baylor, D. A., & Shatz, C. J. (1991). Synchronous bursts of action potentials in ganglion cells of the developing mammalian retina. *Science* **252**: 939–943.

Metzger, W. (1932). Versuch einer gemeinsamen Theorie der Phänomene Fröhlichs und Hazelhoffs und Kritik ihrer Verfahren zur Messung der Empfindungszeit. *Psychologische Forschung* **16**, 176–200.

Mollon, J. D., & Perkins, a. J. (1996). Errors of judgement at Greenwich in 1796. *Nature* **380**(6570): 101–102.

Namba, J., & Baldo, M. V. C. (2004). The modulation of the flash-lag effect by voluntary attention. *Perception* **33**: 621–631.

Nijhawan, R. (1992). Misalignment of contours through the interaction of apparent and real motion systems. *Invest Ophthalmol Vis Sci* **33**(4): 974. See Appendix.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature* **370**(6487): 256–257.

Nijhawan, R. (1997). Visual decomposition of colour through motion extrapolation. *Nature* **386**(6620): 66–69.

Nijhawan, R. (2002). Neural delays, visual motion and the flash-lag effect. *Trends Cogn Sci* **6**(9): 387–393.

Nijhawan, R. (2008a). Visual prediction: psychophysics and neurophysiology of compensation for time delays. *Behav Brain Sci* **31**: 179–239.

Nijhawan, R. (2008b). Predictive perceptions, predictive actions, and beyond. *Behav Brain Sci* **31**(2): 222–239.

Nijhawan, R., & Kirschfeld, K. (2003). Analogous mechanisms compensate for neural delays in the sensory and the motor pathways: evidence from motor flash-lag. *Curr Biol* **13**(9): 749–753.

Nijhawan, R., Watanabe, K., Khurana, B., & Shimojo, S. (2004). Compensation for neural delays in visual-motor behaviour: no evidence for shorter afferent delays for visual motion. *Vis Cogn* **11**: 275–298.

Nijhawan, R., & Wu, S. (2009). Compensating time delays with neural predictions: are predictions sensory or motor? *Philos Trans R Soc Lond A Math Phys Eng Sci* **367**: 1063–1078.

Purushothaman, G., Patel, S. S., Bedell, H. E., & Ogmen, H. (1998). Moving ahead through differential visual latency. *Nature* **396**(6710): 424.

Raiguel, S. E., Lagae, L., Gulyas, B., & Orban, G. A. (1989). Response latencies of visual cells in macaque areas V1, V2 and V5. *Brain Res* **493**(1): 155–159.

Ratliff, F. (1965). *Mach Bands: Quantitative Studies on Neural Networks in the Retina.* San Francisco, CA: Holden-Day.

Rojas-Anaya, H., Thirkettle, M., & Nijhawan, R. (2005). Flash-lag anisotrypy for movement in three domains [Abstract]. *Perception* **34**(ECVP abstract supplement): 219–220.

Roulston, B. W., Self, M. W., & Zeki, S. (2006). Perceptual compression of space through position integration. *Proc R Soc Lond B Biol Sci* **273**(1600): 2507–2512.

Rubin, E. (1929). Kritisches und Experimentelles zur 'Empfindungszeit' Fröhlichs. *Psychologische Forschung* **13**: 101–112.

Schlag, J., Cai, R. H., Dorfman, A., Mohempur, A., & Schlag-Rey, M. (2000). Extrapolating movement without retinal motion. *Nature* **403**: 38–39.

Sheth, B. R., Nijhawan, R., & Shimojo, S. (2000). Changing objects lead briefly flashed ones. *Nat Neurosci* **3**(5): 489–495.

Sillito, A. M., Cudeiro, J., & Jones, H. E. (2006). Always returning: feedback and sensory processing in visual cortex and thalamus. *Trends Neurosci* **29**(6): 307–316.

Sillito, A. M., & Jones, H. E. (2002). Corticothalamic interactions in the transfer of visual information. *Philos Trans R Soc Lond B Biol Sci* **357**(1428): 1739–1752.

Sundberg, K. A., Fallah, M., & Reynolds, J. H. (2006). A motion-dependent distortion of retinotopy in area V4. *Neuron* **49**(3): 447–457.

van de Grind, W. (2002). Physical, neural, and mental timing. *Conscious Cogn* **11**(2): 241–264; discussion 208–213.

Vroomen, J., & de Gelder, B. (2004). Temporal ventriloquism: sound modulates the flash-lag effect. *J Exp Psychol Hum Percept Perform* **30**(3): 513–518.

Whitney, D., & Murakami, I. (1998). Latency difference, not spatial extrapolation. *Nat Neurosci* **1**(8): 656–657.

Wojtach, W. T., Sung, K., Truong, S., & Purves, D. (2008). An empirical explanation of the flash-lag effect. *Proc Nat Acad Sci USA* **105**: 16338–16343.

Part V

Space–time and awareness

28

Object updating: a force for perceptual continuity and scene stability in human vision

JAMES T. ENNS, ALEJANDRO LLERAS, AND CATHLEEN M. MOORE

Summary

How does the visual system provide us with the perception of a continuous and stable world in the face of the spatial–temporal chaos that characterizes its input? In this chapter we summarize several programs of research that all point to a solution we refer to as *object updating*. We use this phrase because perceptual continuity seems to occur at an *object* level (as opposed to an image level or a higher conceptual level) and because our research suggests that the visual system makes a sharp distinction between the formation of new object representations versus the *updating* of existing object representations. We summarize the research that led us to this view in the areas of masking by object substitution, the flash-lag illusion, response priming, and an illusion of perceptual asynchrony.

28.1 Introduction

Biological vision is the marvelous ability of an organism to be informed about its surroundings at a distance and with a high degree of spatial and temporal resolution. This ability allows us to know where things are, what shape and color they are, and equally importantly, *when* they are there, so that we may interact with them appropriately. Yet, contrary to many people's implicit understanding of how biological vision is accomplished, it is *not* a process by which light, reflected from surfaces in the three-dimensional world, is recorded faithfully by the brain in order to reconstruct the nature of the surfaces that gave rise to the recorded pattern of light. To a layperson, this may seem like a plausible place to begin, but modern vision science has pointed to numerous reasons why this approach is a nonstarter. One reason is the computational complexity of the problem of inverse optics (reconstructing the three-dimensional world from a two-dimensional pattern of light on the retina). This has been shown to be an underdetermined problem, in that the time needed to solve it far exceeds the time limits under which biological organisms must act in order to survive (Tsotsos 1990). This forces visual systems to be highly selective in their processing of light.

A second problem with this naïve view is that the pattern of light falling on the retina is highly discontinuous, across both space and time, for many reasons. One is that the retina does not register light in a uniform way across its surface, being, for example, more sensitive

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

to color at the center and more sensitive to motion in the periphery. Second, there is only a relatively small spatial window in which the eye has a high degree of spatial resolution. As such, many eye movements are needed in order to "see" even the simplest of scenes, with the duration and order of these eye movements being extremely difficult to predict (Henderson & Ferreira 2004). But these biological solutions to accomplishing vision (i.e., the need for processing selectivity and discrete eye movements) also present scientists with a new set of problems to understand. For example, how does the visual brain create a sense of timely order out of what seems at first glance to be the spatial and temporal chaos of the visual input to the brain?

In this chapter we highlight one of the general solutions that we believe human vision has settled on in order to establish perceptual continuity and scene stability in the face of discontinuous inputs. We refer to this solution as *object updating* based on our findings that perceptual continuity seems to occur at the *object* level (as opposed to the image level or a higher conceptual level). This solution involves a sharp distinction between the formation of new object representations versus the *updating* of existing object representations. In what follows we present research that led us to this view.

28.2 Object updating

Any visual system that samples information more-or-less continually is faced with a fundamental problem. How should the system incorporate newly sampled information into representations it has formed from past samples? One possibility is that updating occurs via a point-for-point image-comparison process, whereby each lowest-level unit in the representation (retinal cell or pixel) is updated independently. Such a mechanism could be easily implemented as a parallel process, but it would be blind to any meaning in the scene, such as knowing which objects are present, where objects are in relation to each other, and whether any objects are moving.

An alternative possibility is that representations are updated through an object-based process, such that the meaningful units in the scene (objects) are taken into account, and changes are made only insofar as they occur to an object already represented. This distinction between image-based and object-based updating is analogous to differences between pixel-based (painting) and object-based (drawing) programs for computer graphics. In an image-based program, editing something on one object can inadvertently alter another object. For drawing programs, on the other hand, objects are selected and edited independently. Objects other than the one currently selected are protected from changes that are made to the selected object, even if the two objects overlap each other in the image space.

As already noted, image-based updating is appealing from a computational perspective because it can be implemented easily within models that embody the retinotopic registration of information in different visual brain areas. But this appeal must be weighed against the considerable costs it incurs for later, higher-order processing. Because image updating has no regard for the meaning of the scene in terms of surfaces and objects, many important distinctions are lost with each resampling cycle. Image-based updating would fail, for example, to maintain region assignments to figure versus ground, edge assignments to luminance change versus surface orientation, as well as associations between discontinuous regions of a surface caused by occlusion. Given the importance of organized representations for disambiguating the retinal image, it is our view that image-based updating cannot be a complete solution to the problem of perceptual continuity.

This consideration of the difficulties encountered by the image-based approach makes object updating worth considering. The general idea is that, if currently sampled information is interpreted as deriving from an object that is already represented in the scene, then that information will be used to update that object representation. In contrast, if the new information is interpreted as deriving from a different object, then the original object representation will be spared from updating and will therefore remain unchanged in the face of new sensory information. Finally, if the information is perceived as deriving from a new object, then it may elicit the establishment of a new object representation in the scene. Notice that in any given sampling cycle, an old object could be in a new location and a new object could be in a location where an old object had been before. In this way, object-mediated updating is dissociable from image-based updating.

Our ideas on object updating first emerged from research on a type of backward masking that has come to be referred to as masking by *object substitution*, a term introduced by Enns and Di Lollo (1997). However, because the theoretical ideas behind this term are the focus of this chapter, we will begin with a less theory-laden description of the "four-dot masking" that led to these ideas. Backward visual masking, at its most general, refers to the observation that the presentation of a later shape (the *mask*) disrupts the processing of an earlier shape (the *target*) that would have been completely visible if presented in isolation. When we first introduced four-dot masking, the prevailing view of how backward masking is accomplished was that the contours of the mask (a) were confused with the contours of the target in early visual representations (integration), (b) interrupted the ongoing processing of the target contours (interruption), or (c) inhibited the emerging contours of the target (inhibition). However, the details of the four-dot masking procedure caused us to rethink how backward masking might be accomplished.

In a typical four-dot masking experiment (Di Lollo et al. 2000; Enns 2004) a display of shapes is presented (e.g., diamonds with missing corners, circles with gaps in one of four locations, or a variety of simple shapes) and one of the shapes is surrounded by four small dots (sometimes each dot is only one pixel in size). It is the participant's task to identify the shape surrounded by these four dots. When these shapes and the four dots flash on and off simultaneously, it is a relatively easy task for the participant to identify the shape indicated by the dots. However, when only the four dots linger on the screen after the shapes have been erased, the shape surrounded by the dots becomes very difficult (and sometimes even impossible) to identify. It seems as though only the four dots were ever presented in that location (for a demonstration see www.sfu.ca/∼enzo/).

This surprising finding that a sparse pattern of only four dots surrounding the target are effective as a backward mask led us to rethink the prevailing theoretical accounts of backward masking. For one thing, no contour-based interference of any kind, whether resulting from integration, interruption, or inhibition, seemed to be at issue because the contours of the dots were so small in comparison to the contours of the target they were masking, and so great was the distance of the mask contours from those of the target. A traditional understanding of contour interactions in early visual processing predicted that no contour effects should occur under these conditions (Di Lollo et al. 2000). Second, the fact that backward masking occurred even though the mask was visible from the onset of the target display was inconsistent with the then-understood temporal dynamics of masking. Theories of contour interruption and inhibition predicted that masking should be strongest when there was a critical delay between the emerging neural signal of the target and the later arriving neural signal corresponding to the mask (Enns & Di Lollo 2000).

But if contour interactions and temporal asynchronies in the neural signals associated with targets and masks are not responsible for dot masking, then what is? The hypothesis we entertained was that the lingering appearance of the four dots in the target display resulted in the formation of a different object representation for that location in the scene. Initially, the object representation for that location began with the inclusion of the contours associated with both the shape and the four dots. However, before that representation could be completely formed (at least completely enough to lead to a positive identification of the shape) its location in the display contained only the four dots. Thus, the four dots came to be seen by the visual system as an updated version of the object representation that was initiated by the appearance of the target shape and the dots. The continuity of the four dots over time, in comparison to the brief appearance of the target contours without their later confirmation, was taken as evidence by the visual system that only the four dots had ever existed in that location.

In this view of object updating as the mechanism of backward masking, the simultaneous onset of the mask and target elicits an initial grouped representation of their contours. When the target shape disappears before a stable representation of it is formed, but the mask lingers, the scene comes to be represented as the original object changing shape. Information that is then sampled from the lingering display containing only the mask is used to update the original representation. However, because it is the original representation that is needed to answer the question posed by the experimenter (i.e., what is the identity of the target?) this updating process results in poor performance. The relevant information for the psychophysical task is simply no longer included in the updated representation, presumably because this is a design feature of a biological visual system that updates information with a strong bias for object continuity. In keeping with this account, we also noted that attending to the location of the target, either previous to its appearance or rapidly on its presentation, facilitates the establishment of a stable representation (Di Lollo et al. 2000; Enns 2004). If a stable representation of the target can be established based on the brief information about it prior to its offset, then information derived from the lingering mask has less of an opportunity to disrupt critical target information.

The initial evidence that prompted us to think of four-dot masking in terms of object updating was that the mask was so sparse. If only contours were involved, then the target should have had plenty of room to shine through the space in between the widely separated dots in the mask and thereby avoid any masking. But the masking we observed was also not at a purely abstract and conceptual level either. We noted that four dots that were not centered on the location of the target were quite ineffective in masking the target, as for example in the case of four dots that only underlined or lay beside the location of the target (Enns & Di Lollo 1997). Clearly, this form of masking was sensitive to some aspects of space, implying that it was not occurring because of a more general interruption of processing. However, its space sensitivity was also not as precise as individual contours either, because masking was quite effective when the four masking dots that began in the same location as the target lingered in an adjacent location eccentric to the center of gaze (Jiang & Chun 2001). Another clue to the object-based nature of the masking was that the strength of masking increased with the number of potential target objects in the display (Di Lollo et al. 2004). Masking strength also decreased if the location of the target or the identity of the target was cued in advance (Enns 2004), in keeping with the position that there is a limited capacity to represent objects in visual short-term memory (Vogel et al. 2001).

Lleras and Moore (2003) provided the first direct test of the hypothesis that four-dot masking was object-mediated. Their experiments were designed to test whether an object-level component of four-dot masking could be isolated from any spatially local interference that might also be contributing to it. Recall that four-dot masking occurs when the mask lingers on the screen for some time following the offset of the target. Under the object-updating account, the important aspect of the delay is that the object that corresponds to the target is delayed in the emerging mental representation of the scene, not that a masking object is physically delayed at the same location as the target. With this in mind, Lleras and Moore used apparent motion to delay the mask as an object in the scene without delaying it at the location where the target object appeared. In the critical condition, a given mask-target pair offset at the same time, but a short while later, the mask reappeared at a new location. This gave rise to the perception of a single object, moving and changing in shape over time. Thus, at the level of object representations of the dynamic scene, this was a delayed-offset condition because the object that started as the mask-target pair and changed to just the mask was delayed in the scene. At the local or image level, however, it was a simultaneous-offset condition because the mask and target offset at the same time at the same location. Yet, despite this simultaneous offset, substantial masking occurred.

In contrast, no significant masking occurred in a long-interval condition in which the same transient events as those in the short-interval condition occurred, but the interval was now too long to support apparent motion. What happened here instead was that the display was perceived as one object suddenly appearing at one location, turning off, and then a new object appearing at a new location. In the strongest instantiation of this research logic, the mask was only a single dot that never appeared at the location of the target, and yet it still produced significant masking (Lleras & Moore 2003, Experiment 4).

Moore and Lleras (2005) reported additional direct evidence that four-dot masking involved object-mediated updating. They demonstrated that targets could be protected from four-dot masking by introducing manipulations that facilitate the early establishment

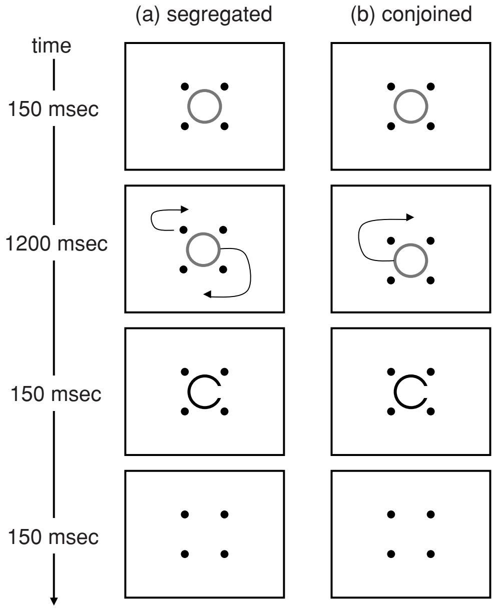


Fig. 28.1 An illustration of the display sequence for one of the items in Moore and Lleras (2005). (a) In the segregated condition the four dots and the circle moved independently of one another before the masking sequence. (b) In the conjoined condition the four dots and the circle moved in concert with one another.

of separate representations for the target and the mask. In other words, if the visual system treats the mask as a different object from the target, then the information sampled from a lingering mask will not be used to update the target representation, and whatever target information was established will be spared from the updating process.

Figure 28.1 illustrates an example of how this logic was applied, using segregation by common motion to manipulate the degree to which masks and targets were represented as single or separate objects. The figure shows only one of the many items in each display. The participant's task was to find the one dark circle-with-a-gap among the many light circles-with-gaps and to report the direction in which that target circle's gap pointed (up, down, left, right). Each trial began, however, with placeholders that were circles surrounded by four dots. There were no gaps in the placeholder circles, and they were all the identical shade of gray. Therefore there was no indication as to which item would eventually be the target, and no information about the identity (i.e., gap direction) of any of the circles. The displays did, however, allow for the establishment of initial object representations for the circles and dot masks. Before displaying the target information, a short movie was shown. In the *segregated* condition, the movie was of the circles and masks jiggling around their center point independently. In the *conjoined* condition, the circles and masks jiggled around their center point together as a unit. The movie ended in both conditions, with the circles and masks recentered within each pair, and the trial then unfolded as a usual dot-masking trial. Thus, the presentation of the target information was identical in the two conditions; they differed only in the object history with which the viewer approached that information. Consistent with the object-updating hypothesis, that history influenced dot masking, such that masking was eliminated in the *segregated*, but not in the *conjoined,* condition. Our interpretation is that by establishing an object representation for the target distinct from the mask early in processing, the target was protected from updating, and therefore from four-dot masking. Several different strategies of facilitating the establishment of separate object representations for the target and mask early in processing render the same effect (Moore & Lleras 2005).

Together, these studies provide a dissociation between the effects of image location and object representation in four-dot masking that clearly implicates the object as the vehicle of interference. In particular, as long as the mask is represented as the same object as the target, interference can occur even when the mask never appears at the location of the target (Lleras & Moore 2003). At the same time, if the mask is represented as a different object from the target, the interference does not occur even when the mask lingers at the location of the target (Moore & Lleras 2005).

28.3 Object updating and the flash-lag illusion

The success we had in understanding backward masking in terms of object updating prompted us to examine several other visual phenomena in which errors that participants make are consistent with the effects of having lost earlier information about an object in favor of updated information. We refer to this class of errors as those of *perceptual lag*, because the incorrect responses given by participants are not random guesses. Rather, they can be linked directly to a competition between features that were presented first for some object versus features that were presented more recently, with the errors showing a strong bias for the features presented most recently.

An interesting phenomenon of this kind is the *flash-lag illusion* that involves a misperception of the spatial relations between a moving object and a briefly flashed one. This illusion is of particular interest to us because it seems to illustrate the object-updating process in one of the most basic of perceptual experiences, the seen position of an object in space while it is undergoing motion.

A typical version of the illusion is illustrated in Fig. 28.2. A disc travels in a circular path and at some point in its journey a square is flashed briefly. If the disc continues on its path, past the frame in which it was aligned on a radial axis with the square, then the moving disc tends to be seen as spatially "ahead" of the briefly flashed stationary square, even though when the square appeared, the objects were aligned (Mackay 1958; Nijhawan 1994). But if instead of continuing on its motion path, the moving disc is erased immediately after it is aligned with the square, then the disc and the square are seen accurately as aligned (Whitney & Murakami 1998; Eagleman & Sejnowski 2000; Whitney et al. 2000).

Moore and Enns (2004) examined whether the different perceptions associated with these two conditions, which they referred to as *continued-motion* versus *stopped-motion*, could

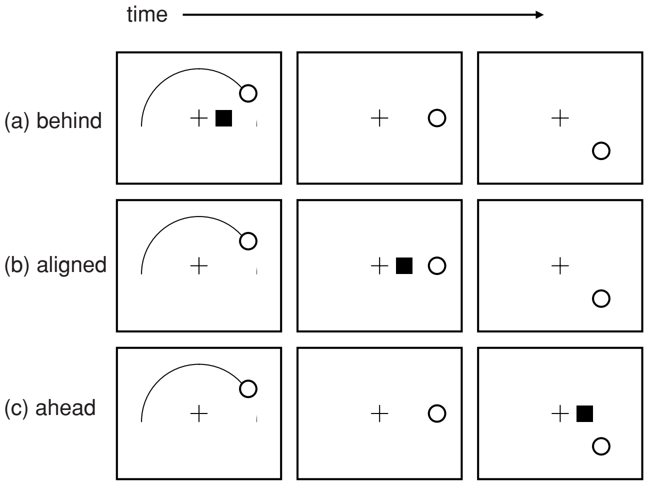


Fig. 28.2 Illustration of the displays used in Moore and Enns (2004). A moving disk (circle) moved around the circumference of an imaginary circle. At some point in this journey a flash occurred (square), and it was the participant's task to indicate whether the moving circle was (a) behind, (b) aligned, or (c) ahead of the square at the moment it flashed.

be understood in terms of object updating. (Other authors in this volume refer to these conditions as "continuous cycle" and "flash-terminated," respectively). Our guiding hypothesis was that the difference in perception between these two conditions hinges directly on the object-updating process. When the disc moves to a new position following the flash in the continued-motion condition, the new position information replaces that acquired at the time of the flash. In contrast, when there is no new position following the flash in the stopped-motion condition, there is no new position information to update the previous information, and so the alignment of the two objects is perceived accurately. Demonstrations of the conditions tested in this study can be found online: www.psych.ubc.ca/∼ennslab/research.

Moore and Enns (2004) provided direct evidence for this interpretation of the flash-lag illusion in three steps. In a first phase of the study we confirmed that the continued- and stopped-motion conditions indeed led to a very different set of perceptual reports. A large flash-lag illusion was recorded in the continued-motion condition and no illusion was observed in the stopped-motion condition. In the second phase, changes were made to the displays such that on some trials participants saw the moving disc as undergoing an abrupt and large change in either size or color at the moment of the flash. These changes were intended to disrupt the normal process of object continuity. Our reasoning was that if an object was seen to change in a radical way then a new object representation would have to be formed for it after the change occurred. Consistent with our prediction, under these conditions the flash-lag illusion was no longer experienced by participants, even though the motion was as continuous as it was before. The change in size or color of the moving disc was enough to disrupt the experience of perceptual continuity for that object.

In phase three of the study we tested a unique prediction of the object-updating account of the flash-lag effect. If the disruption in perceptual continuity caused by the large changes in size or color actually leads to the formation of a new object representation, then we should be able to find direct evidence for this new representation in the form of a change in the total number of objects perceived in the no-change (standard flash-lag) versus change (disrupted size or color) conditions. Specifically, two discs should be visible at the moment of the flash in the changed condition but not in the unchanged condition. This was tested by asking participants to report on the number of discs that were visible at the time of the flash. The possibility of illusory reports of more than one disc was anchored to reality by including trials in which there actually were two discs presented. The results showed that the change in size or color of the moving disc led to the perception of "double discs," with one corresponding to the original moving disc and the other corresponding to the suddenly changed disc. Note that this is consistent with the flash-lag illusion actually being a form of backward masking, as described in the previous section, where the spatial position of the disc seen after the flashed square is automatically and irretrievably updated so that the actual position at the time of the flash is lost. In the case of the large change in size or color, the updated spatial position with regard to the original disc is not assigned to the suddenly changed disc. Thus, the changed disc is spared from the normal process of object updating and its spatial position can be perceived accurately.

In conclusion to this section, we want to note that our object-updating account is not at odds with previous accounts of flash-lag illusion, which have focused either on visual integration over a temporal window (Eagleman & Sejnowski 2000; Krekelberg & Lappe 2000), visual extrapolation into the future (Nijhawan 1994, 2002; Schlag & Schlag-Rey 2002), or the relative speed of neural signals from moving and static stimuli (Whitney & Murakami 1998; Whitney et al. 2000; Whitney 2002). Those theories have all offered various mechanisms to account for the illusory percept and yet they have remained silent on the level of representation that is involved in the illusion. As such, we propose our object-updating account of the illusion to add a critical level of detail to the mechanisms involved in each of these theories. The object-updating account contributes to an understanding of the illusion at a level of analysis that has so far been ignored, namely, the critical role of perceptual objects in flash-lag sequences.

28.4 Object updating and response priming

So far we have discussed how object updating influences the representations that form the basis of our conscious awareness. However, we do not only use vision to help us identify objects, we also use vision to help us interact with those objects through actions. That is, we sometimes grasp objects, catch them, navigate around them, or simply point to them. There is now a large body of evidence that visual information is processed by two distinct and somewhat independent neurological subsystems in the primate cortex: a ventral pathway that is specialized for the conscious perception of objects and scenes, and a dorsal pathway that is specialized for visually guided action (Milner & Goodale 1995; Goodale &

Humphreys 1998). From this perspective, our previous findings on object updating can be characterized as vision solely for conscious perception. We therefore began to ask whether object updating is also relevant for the representations that inform the subsystem responsible for visually guided action.

Masked priming is a behavioral tool used by researchers to better understand how visual processing influences motor responses. The method is simple, consisting of the presentation of three shapes in rapid succession: a prime, followed by a mask, which is itself followed by a target. The participant is instructed to respond to the target as rapidly as possible without making too many errors, usually by making a left or a right key press. The purpose of the prime is to provide advance information about the upcoming target and therefore the response that will be required. The typical finding is that when the prime specifies the same response as the target (compatible trial), then response time to the target is faster and more accurate than when the prime specifies the opposite response from the target (incompatible trial). This occurs even under conditions in which the prime shape is not visible. This result is taken as evidence that the motor system starts preparing its response even before conscious perception occurs, and the magnitude of this preparation is indexed by the difference between incompatible and compatible responses.

The role of the mask in traditional studies of masked priming has been to limit the strength or duration of the perceptual representation of the prime. This is done by varying the period of time that the prime may be processed prior to the appearance of the mask and/or by varying the intensity of the mask (the number and contrast of its contours along with its duration). In other words, masking is used as a tool of convenience to limit the exposure of the visual system to the prime shape. However, this intuitive and widely accepted use of masked priming fails to acknowledge that mask shapes do not merely decrease the visibility of a prime shape. In fact, as we have seen in previous sections, masks can fundamentally alter the way in which a prime is processed.

Lleras and Enns (2004) provided a striking demonstration of just how much the choice of a mask can alter the processes under investigation in masked priming. We began this study when we learned about the so-called negative compatibility effect (NCE), first described by Eimer & Schlaghecken (1998). As suggested by the word *negative* in the name, the NCE is a masked priming effect in which *compatible* primes actually lead to slower and less accurate responses than *incompatible* primes. This counterintuitive finding has been studied extensively, and several theories have been put forward, but what struck us was that the masks used in previous studies shared a commonality: the mask was composed of the same visual features that were used to create the prime and target shapes. For example, when the prime and target were each comprised of arrows (pointing either left or right), the mask consisted of a single image obtained by superimposing the same left- and right-pointing arrows. When these masks followed the primes, thereby reducing their visibility, negative priming was obtained. However, when these masks were omitted, and the primes were perfectly visible, then the positive priming effect obtained in most previous studies of masked priming was observed (Eimer & Schlaghecken 1998; Klapp & Hinkley 2002).

To be sure, the shared similarity between the primes and these masks suggested that these masks should be very effective in their reduction of the participant's ability to respond to the primes when specifically asked to do so, which they were. But we suspected that these masks might be doing something else besides simply limiting the visibility of the prime shapes.

Our past experience with the concepts of object updating led us to test a very different hypothesis regarding the critical ingredient for the NCE (Lleras & Enns 2004). It was prompted by our observation that the display sequences used in the NCE were similar to those in which we had previously studied object updating in masking and the flash-lag effect. Specifically, prime and mask shapes were presented close together in both time (intervals between separate shapes were on the order of 50–150 msec) and space (shapes were presented at the same or in nearby locations). Moreover, the shared features among the prime-mask-target shapes might readily induce the visual system to interpret these sequences not as discrete and formally unrelated events (as they were interpreted by the experimenters) but as instantiations of the same object changing rapidly over time. If so, even though the mask shape might be formally neutral, in terms of the information it provided with regard to the upcoming target and its required response, the mask might be informative to a visual-motor system that was rapidly incorporating new information about an emerging object representation.

To illustrate our thinking, consider the case of a right-pointing arrow prime followed by a mask consisting of superimposed left and right arrows, and this mask itself being followed by a right-pointing target. When the right-pointing prime first appears, the visual system may begin to form an object representation (let's call it P), such that the attributes of the prime will be encoded and linked to P. If some of these prime attributes are strongly associated with motor responses, even in the very early stages of their formation, this representation will get linked to its associated response and give rise to corresponding motor preparation. In this case, preparation for a right response will begin. When the mask shape appears, especially if it shares task-relevant features with the prime, it is likely to be interpreted as a new and updated instantiation of P. Therefore, the mask will not receive its own object representation but will instead be incorporated into the existing representation of P that has already been initiated. This means that task-relevant information now detected in the mask that was not already present in P will be added to it. In this example, the new features correspond to a left-pointing arrow, the right-pointing features having already been encoded.

This kind of thinking led to our hypothesis that in the NCE, priming is determined by the most recently detected set of response-relevant features in a representation that has been updated as part of the ongoing task of the participant to respond rapidly to the target object. We tested this hypothesis in a series of experiments by comparing the priming effects obtained from masks comprised of either task-relevant or task-irrelevant visual features. Fig. 28.3(a) shows a typical sequence of trial events and Fig. 28.3(b) shows a summary of the priming effects observed under the masking conditions we tested. As can readily

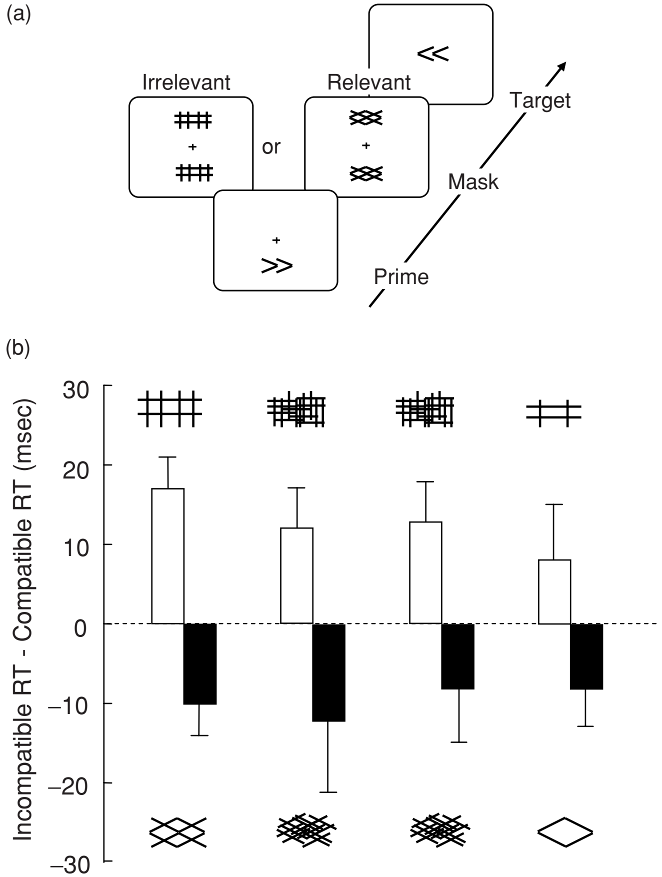


Fig. 28.3 (a) A typical sequence of trial events in an experiment on the negative compatibility effect. (b) Mean priming effects (incompatible minus compatible response time) as a function of various masks inserted between the prime and the target arrows in Lleras and Enns (2004). The specific masks used are shown above and below each data point. Positive priming effects (white bars) occurred for masks that had no features relevant to the target discrimination task. Negative priming effects (black bars) occurred when masks contained features that were relevant to the target discrimination task.

be seen, masks consisting of task-relevant features invariably produced negative priming effects, replicating previous studies of the NCE, whereas irrelevant masks systematically yielded positive priming effects, as expected based on our views of object updating.

These results directly refute theories of masked priming in which the critical ingredient is the strength of the representation corresponding to the prime, which is usually assessed by measuring the visibility of the prime shapes. In such theories, all masks are equal provided that they reduce the visibility of the prime sufficiently to activate the unconscious processes associated with the subsequent target. However, this was not what we found. Rather, we found that completely opposite priming effects could be obtained at equal levels of high- or low-prime visibility. What determined the direction of the priming in both cases was whether the mask contained task-relevant (negative priming) or task-irrelevant features (positive priming).

The alternative theory we proposed to account for the NCE was based on our interpretation of backward masking and the flash-lag, namely, that the prime and mask may be interpreted by the visual system as a change in the same object over time. As a result, the prime–mask sequence is susceptible to the normal processes of object updating, and a single representation may represent both stimuli as one object changing over time. According to this interpretation, although the observed priming effect may superficially appear to be negative with respect to the prime, it actually corresponds to a positive priming effect induced by the novel task-relevant features in the mask. In sum, the observed priming effect reflects the influence not of the prime alone, but of the prime–mask bundle, on response selection. On irrelevant-mask trials, the most recent set of task-relevant features encoded prior to target onset are those present in the prime because the mask has none of these features. In this case then, the observed priming effect really does reflect the positive influence of the prime on response selection to the target.

Lleras and Enns (2004) also showed that this object-updating account could be applied to masks that did not contain exact replicas of the primes but also to randomly oriented line masks that merely contain similar features in spatially noncorresponding locations to those in the target. Once again, this is consistent with an object-based account of updating, in that what is updated is a representation of a group of features and not merely an updating at the level of specific contours or surface features.

In summary, these experiments help emphasize that the behavioral consequences of masking should always be examined directly rather than simply being assumed. They also demonstrate that object updating is a powerful concept that can be applied not only to the understanding of conscious perception but also to unconscious processes that lead to motor responses. Finally, they show that the concept of object updating leads to testable predictions and that it can help explain some behavioral phenomena with a simpler set of assumptions than those needed when masking is used as a tool to merely reduce visibility.

28.5 Object updating and the perceptual asynchrony illusion

When participants view a display in which all objects alternate for equal durations between two colors and two directions of motion, the apparent coincidence of a specific color and motion does not always match their physical coincidence. This is the *perceptual asynchrony illusion* (PAI), and we have recently begun to examine the possible role played by object updating in the illusion (Oriet & Enns, under review).

In the seminal study of the PAI (Moutoussis & Zeki 1997a), participants viewed numerous squares that each alternated between red and green while at the same time alternating between moving up and down. Participants were asked to report the color of the squares while they were moving upward. Different feature durations were tested (e.g., 250 msec and 350 msec) along with different phase relationships between the changes in color versus motion direction (e.g., simultaneous or 0 deg changes versus changes that were 90, 180, or 270 deg out of phase). The duration of each feature was not nearly as influential in the color reports of participants as was temporal asynchrony. When the change in color preceded the change in motion (e.g., upward moving squares were initially red and then turned green), the reported color was the one presented during the last portion of the motion period (e.g., upward squares were reported as green). In fact, changes in motion direction had to precede changes in color by 80 msec before changes in the two features were reported as coincident.

The PAI has attracted considerable theoretical interest because it has been claimed as a way to index the modularity of consciousness (Moutoussis & Zeki 1997a, 1997b; Zeki & Bartels 1998) or at least the relative time required for different types of visual properties to become accessible for conscious report (Arnold et al. 2001; Arnold & Clifford 2002; Nishida & Johnston 2002). In contrast, our interest in the PAI was piqued when we saw that almost no consideration had been given to the possibilities that (a) the illusion was influenced by the need to switch attention from one feature to another, or (b) the processes of object updating may be playing a role in this illusion as well. The fact that both of these possibilities had been overlooked was surprising to us because the PAI bears at least superficial resemblance to the other visual phenomena we have reviewed, such as backward masking, the flash-lag effect, and masked priming. In each of these effects, the focus of attention plays an important role and each of them involves errors of perceptual lag.

To be more specific, in each of the effects we have already discussed (i.e., masking, flash-lag, masked priming) there is always a *defining feature* that participants must become aware of first, before they are required to turn their attention to the *report feature*. For backward masking, the defining feature is the detection of a visible mask shape, which is the cue to try to report the immediately preceding target shape; in the flash-lag the defining feature is the flashed object, which is the cue to report the spatial position of the moving object. In the case of the PAI, participants must first see a defining direction of motion in a square before they are able to report on the color of the same square. If this similarity between tasks is more than superficial, then factors that influence the speed or difficulty of this switch in attention from the defining to the report feature should also influence the PAI. We therefore thought it was worth checking to see whether this illusion was influenced by the tendencies of the visual system to be somewhat sluggish in switching attention from one visual feature to another, and in the dynamic updating of object representations as new information becomes available.

In a first experiment, Oriet and Enns (under review) tested whether (1) perceptual lag errors are greater when features change asynchronously rather than synchronously, and (2) errors in either of these types of display are influenced by the difficulty of feature detection. Displays consisted of a series of 16–20 moving checkerboards that varied in both color and motion direction. The participant's task was to detect the single red checkerboard in the series (defining feature = color red) and to report the direction of motion taken by the red checkerboard (report feature = one of four motion directions). Having only a single target (red checkerboard) in the series allowed us to determine whether errors in reports of motion direction reflected a perceptual lag, in which case errors would come from the motion directions that occurred *after* the target color, or whether they merely reflected a noisy signal, in which case errors should distribute randomly from values occurring before and after the target color.

The results did not support the hypothesis that slow attention plays the most important role in the PAI. To be sure, lag errors were more likely when displays changed more rapidly than when they changed slowly. This confirms that when participants err in their report of the features in a rapid sequence, they are not simply guessing but are responding based on features they see later in the sequence. However, the most important finding in this experiment was that lag errors occurred more frequently for asynchronous than for synchronous displays at the same display rates. If the time required to switch attention from defining to report features was the primary limiting factor in the illusion, then these two types of displays should show similar lag errors at the same rates of display change. This means that the PAI arises from factors that go well beyond a simple sluggishness in being able to shift attention from one feature to another.

The importance of temporal synchrony for feature binding has been documented previously. Perceptual objects can be defined solely by the common behavior of their parts in the temporal domain (Usher & Donnelly 1998; Lee & Blake 1999; Sekuler & Bennett 2001). The unusual circumstance that confronts the visual system in the PAI is that features are beginning and ending out of temporal step with one another. This likely prompts the system to look for other clues to how the scene might be organized, such as grouping by spatial proximity and other geometric heuristics relevant to object perception. And while this is going on, the scene continues to change, so the system becomes vulnerable to the feature updating and overwriting that is normally beneficial as one views what first might look like a bird, but then gets reinterpreted as an airplane or even Superman.

To explore an object-updating interpretation of the PAI, Enns and Oriet (under review) conducted three more experiments. In one, they reasoned that if the PAI involves a fundamental ambiguity about the temporal characteristics of features that otherwise appear to be present in the same object, then it should be possible to reverse the direction of the illusion by switching the role of defining and report features. In other words, the PAI should reverse when participants report the color of the upward moving squares versus the motion direction of the red squares. When the roles of defining and report features were reversed in this way, for two different groups of participants viewing the same display sequences, the main finding was that perceptual lag errors were linked to the report feature, not to whether the feature was color or motion.

In another experiment, it was reasoned that the principles of object updating should apply equally well when the changing features are different sensory dimensions (e.g., color and motion direction, as in many previous studies) as when they are from the same sensory dimension (e.g., both are color features). They also tested whether the illusion was stronger when the features that needed to be linked in time were associated with the same object versus when they were associated with separate objects. They reasoned that the object-updating processes that lead to perceptual lag errors would be stronger when the changing features came from the same object, as has already been observed in the flash-lag illusion (Moore & Enns 2004). Both of these hypotheses were confirmed.

In a last experiment, Enns and Oriet (under review) tested whether the illusion would vary with the perceptual grouping of features based on spatial proximity. According to the object-updating hypothesis, the PAI should occur with greater likelihood when the changing features occur in the same location in space (promoting object updating) versus when they occur in different locations (preventing object updating). This hypothesis was also confirmed.

Taken together, we interpret the results of the Enns and Oriet (under review) study as strong support for the idea that the PAI is a consequence of the ubiquitous perceptual updating process that occurs when objects are viewed in the context of dynamic and changing visual scenes. We believe that under normal circumstances, the temporal onset and offset of feature values are concurrent when they derive from the same object. The dark color and characteristic motion pattern of a bird both begin and end with the appearance and disappearance of the bird from view.

28.6 Conclusion

We began this chapter by asking how the visual brain creates a sense of order out of the seeming chaos of the visual input, so that it can arrive at the perception of a stable world, yet one in which objects may also reasonably change their characteristics over time. We proposed one very general solution to this problem based on the idea that visual representations are mediated at the level of perceived objects, rather than at the level of image features or at even higher conceptual levels of representation. We reviewed work demonstrating the consequences of object-mediated updating across a wide range of visual phenomena, including backward masking, the flash-lag illusion, unconscious response priming, and the perceptual asynchrony illusion. The ease with which a single theoretical construct – object updating – can help account for such a wide range of phenomena gives us confidence in the account. At the same time, we think it points to the generality of the process. Object updating is effective in providing stability for perception in the face of highly unstable sensory input and it appears to be ubiquitous in the visual processing of dynamically changing scenes.

Although evidence for object updating is robust in the laboratory studies we have described, we must hasten to add that the conditions in these studies were tailored to give insight into the emerging or online development of perceptual representations. The conditions were specifically designed to show that visual information obtained later in time could influence the emerging representation of an object. From this it should not be concluded that all perceptual representations formed from brief glimpses are necessarily susceptible to these forces. That is, we are not implying that there are no stable representations or that all perceptions are vulnerable to the updating process. Indeed, an important next goal in our research is to better understand what underlies the difference between representations that are vulnerable to updating versus those that are immune. We suspect that visual attention will play a key role in this distinction, and that when we understand its role we will be better positioned to link current understanding of visual perception with that of visual short-term memory.

Acknowledgments

JTE was funded by a Discovery Grant from NSERC (Canada), AL by NSF (Awards 0527361 and 0309998), and CMM by NIH (MH067793).

References

Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: the psychophysics of reentrant visual processes. *J Exp Psychol Gen* **129**: 481–507.

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science* **287**: 2036–2038.

Eimer, M., & Schlaghecken, F. (1998). Effects of masked stimuli on motor activation: behavioral and electrophysiological evidence. *J Exp Psychol Hum Percept Perform* **24**: 1737–1747.

Enns, J. T. (2004). Object substitution and its relation to other forms of visual masking. *Vision Res* **44**: 1321–1331.

Enns, J. T., & Di Lollo, V. (1997). Object substitution: a new form of masking in unattended visual locations. *Psychol Sci* **8**: 135–139.

Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends Cogn Sci* **4**: 345–352.

Goodale, M., & Humphreys, G. K. (1998). The objects of action and perception. *Cognition* **67**: 181–207.

Henderson, J. M., & Ferreira, F. (2004). *The Interface of Language, Vision, and Action.* New York: Psychology Press.

Jiang, Y. H., & Chun, M. M. (2001). Asymmetric object substitution masking. *J Exp Psychol Hum Percept Perform* **27**(4): 895–918.

Klapp, S. T., & Hinkley, L. B. (2002). The negative compatibility effect: unconscious inhibition influences reaction time and response selection. *J Exp Psychol Gen* **131**: 255–269.

Krekelberg, B., & Lappe, M. (2000). A model of the perceived relative positions of moving objects based upon a slow averaging process. *Vision Res* **40**: 201–215.

Lee, S. H., & Blake, R. (1999). Visual form created solely from temporal structure. *Science* **284**: 1165–1168.

Lleras, A., & Enns, J. T. (2004). Negative compatibility or object updating? A cautionary tale of mask-dependent priming. *J Exp Psychol Gen* **133**: 475–493.

Lleras, A., & Moore, C. M. (2003). When the target becomes the mask: using apparent motion to isolate the object-level component of object substitution masking. *J Exp Psychol Hum Percept Perform* **29**: 106–120.

MacKay, D. M. (1958). Perceptual stability of a stroboscopically lit visual field containing self-luminous objects. *Nature* **181**: 507–508.

Milner, A. D., & Goodale, M. A. (1995). *The Visual Brain in Action.* London: Oxford University Press.

Moore, C. M., & Enns, J. T. (2004). Object updating and the flash-lag effect. *Psychol Sci* **15**: 866–871.

Moore, C. M., & Lleras, A. (2005). On the role of object representations in substitution masking. *J Exp Psychol Hum Percept Perform* **31**: 1171–1180.

Moutoussis, K., & Zeki, S. (1997a). A direct demonstration of perceptual asynchrony in vision. *Proc R Soc Lond B Biol Sci* **264**: 393–399.

Moutoussis, K., & Zeki, S. (1997b). Functional segregation and temporal hierarchy of the visual perceptive systems. *Proc R Soc Lond B Biol Sci* **264**: 1407–1414.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature* **370**: 256–257.

Nijhawan, R. (2002). Neural delays, visual motion and the flash-lag effect. *Trends Cogn Sci* **6**: 387–393.

Oriet, C., & Enns, J. T. (under review). The perceptual asynchrony illusion: object updating and unbinding take time. *J Exp Psychol Hum Percept Perform.*

Schlag, J., & Schlag-Rey, M. (2002). Through the eye, slowly: delays and localization errors in the visual system. *Nat Rev Neurosci* **3**: 191–200.

Sekuler, A. B., & Bennett, P. J. (2001). Generalized common fate: grouping by common luminance changes. *Psychol Sci* **12**: 437–444.

Tsotsos, J. (1990). Analyzing vision at the complexity level, *Behav Brain Sci* **13**: 423–445.

Usher, M., & Donnelly, N. (1998). Visual synchrony affects binding and segmentation in perception. *Nature* **394**: 179–182.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *J Exp Psychol Hum Percept Perform* **27**: 92–114.

Whitney, D. (2002). The influence of visual motion on perceived position. *Trends Cogn Sci* **6**: 211–216.

Whitney, D., & Murakami, I. (1998). Latency difference, not spatial extrapolation. *Nat Neurosci* **1**: 656–657.

Whitney, D., Murakami, I., & Cavanagh, P. (2000). Illusory spatial offset of a flash relative to a moving stimulus is caused by differential latencies for moving and flashed stimuli. *Vision Res* **40**: 137–149.

Zeki, S., & Bartels, A. (1998). The asynchrony of consciousness. *Proc R Soc Lond B Biol Sci* **265**: 1583–1585.

29

A motion illusion reveals the temporally discrete nature of visual awareness

RUFIN VANRULLEN, LEILA REDDY, AND CHRISTOF KOCH

Summary

Quasi-periodic or "discrete" brain processes are, in theory, susceptible to a phenomenon known in engineering as "temporal aliasing." When the rate of occurrence of events in the world is fast enough, the perceived direction of these events may be reversed. We have recently demonstrated that, because of a quasi-periodic attentional capture of motion information, continuously moving objects are sometimes perceived to move in the wrong direction (the "continuous Wagon Wheel Illusion"). Using a simple Fourier energy model of motion perception, we established that this type of attentional capture occurs at a rate of about 13 Hz. We verified with EEG recordings that the electrophysiological correlates of this illusion are restricted to a specific frequency band around 13 Hz, over right parietal regions – known for their involvement in directing attention to temporal events. We summarize these results and discuss their implications for visual attention and awareness.

29.1 Introduction

With respect to the temporal organization of visual perception – the topic of this book – one important issue that has puzzled scientists for more than a century (James 1890; Pitts & McCulloch 1947; Stroud 1956; White 1963; Shallice 1964; Harter 1967; Varela et al. 1981; Purves et al. 1996; Crick & Koch 2003; VanRullen & Koch 2003) is whether our experience relies on a continuous sampling or a discrete sequence of periodic "snapshots" or "perceptual frames" of the external world. Although it may seem that such radically different mechanisms should be easy to distinguish using elementary introspection, the realism of the cinema serves to remind us that these two alternatives can in fact lead to equivalent perceptual outcomes. The question, thus, is not a trivial one.

One way to tackle it is to use the fact that quasi-periodic or discrete processes are susceptible to a phenomenon known in engineering as "temporal aliasing," whereas continuous processes are not. Temporal aliasing means that a periodic stimulus can be incorrectly perceived by a system if its temporal frequency is fast enough compared to the system's sampling frequency. Recently, one such temporal aliasing effect has been uncovered in visual motion perception (Purves et al. 1996): a periodically moving stimulus (e.g., a rotating wheel) is perceived, at certain temporal frequencies, to move in the opposite direction, even

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

while it is illuminated by constant light (Schouten 1967). Purves and colleagues reported that this phenomenon occurs for rotational as well as translational motion at temporal frequencies between 2 and 20 Hz, and cannot be explained by spurious eye movements (Purves et al. 1996). This effect, the "continuous Wagon Wheel Illusion," (c-WWI) constitutes a unique tool in the study of the temporal organization of visual motion perception. Here we review our recent investigations of the c-WWI effect, its electrophysiological correlates, and its potential implications for the discrete nature of visual awareness.

29.2 Psychophysical evidence

In our initial experiments we attempted to characterize the properties of the c-WWI effect in terms of spatial and temporal frequency selectivity (VanRullen et al. 2005). For this purpose, we used a custom-made apparatus composed of a motor connected to a speed controller in a closed-loop system, which allowed for real-time readout and adjustment of the motor rotation frequency. Cardboard disks with radial "wheel" patterns were attached to the shaft of the motor. By changing the number of spokes on these wheels or by adjusting the rotation frequency of the motor, we could independently manipulate the spatial and temporal frequency of the stimulus, respectively, and estimate the effect of these manipulations on the c-WWI. Following up on the results of Kline and collaborators (Kline et al. 2004) as well as our own preliminary observations, we treated this illusion as a bistable percept, in which the correct and erroneous directions of motion constantly compete to dominate the subject's perception. Therefore, we asked our subjects to report continuously throughout the duration of 1-minute-long trials the perceived direction of motion (by pressing the corresponding arrow key on a computer keyboard). We estimated the strength of the c-WWI effect as the percentage of trial time spent reporting the illusory direction of motion (a conventional measure of bistable perception). The effects on the c-WWI of varying temporal frequency for one given wheel are illustrated in Fig. 29.1(a). A clear increase in the illusion strength is apparent for temporal frequencies around 10 Hz. This was also the case for two other wheels, spanning a range of spatial frequencies of two octaves overall (not shown here). The fact that the illusion depends mostly on temporal frequency, but not on spatial frequency or the velocity of the moving stimulus (as found also by Simpson et al. 2005) suggests that this effect can be attributed to temporal aliasing, and may reveal discrete processes involved in motion perception.

To pursue these investigations with various types of well-controlled motion stimuli, it is essential to move from a constraining "real-world" stimulus to more controllable, computer-generated ones. In order to do so, the main consideration is that the refresh rate of the display system should be fast enough to avoid it generating its own temporal aliasing artifacts (Burr et al. 1986). In practice, this implies that the refresh rate should be at least twice as high as the fastest temporal frequency to be displayed (this is called the "Nyquist" frequency of the stimulus). We verified that our previous results with real wheels physically rotating at frequencies up to 40 Hz (Fig. 29.1(a)) could be replicated on a computer monitor refreshed at 120 Hz. Indeed, the results were very similar to the previous ones, with a clear

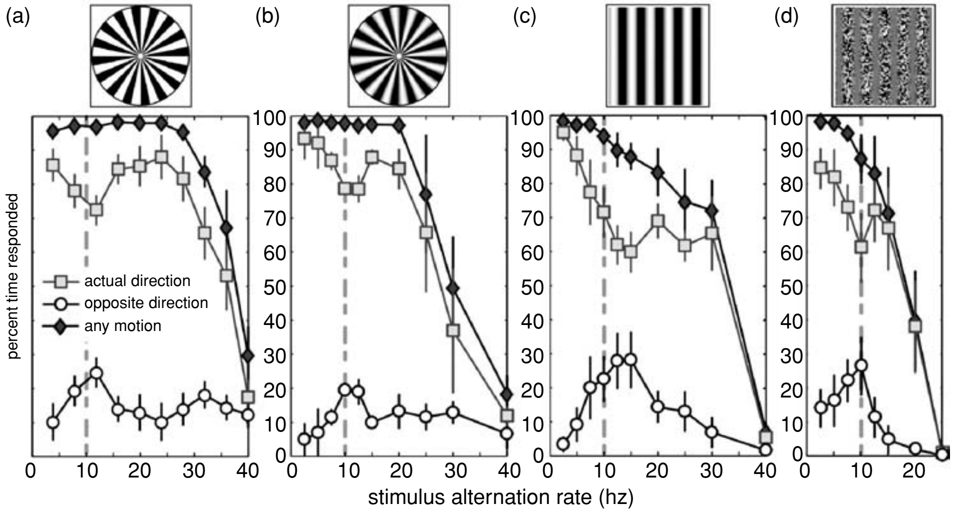


Fig. 29.1 Moving stimuli were displayed continuously for 60 sec while subjects reported the perceived direction of motion using computer keyboard arrows. Although the dominant direction of motion was always the actual direction, observers also reported perceiving the opposite direction for a considerable fraction of the time. Whether using (a) rotating sunburst patterns under natural, continuous illumination ($n = 6$), (b) rotating radial sinusoidal patterns ($n = 4$), (c) horizontally drifting luminance-modulated (first-order motion), or (d) contrast-modulated gratings (second-order motion) on a computer monitor with a 120- or 160-Hz refresh rate ($n = 5$), the c-WWI effect always peaked for alternation rates around 10 Hz. This is indicative of a temporally specific mechanism underlying the illusion.

peak for the c-WWI effect at 10 Hz (Fig. 29.1(b)). In all the following experiments, we used a computer monitor with a refresh rate of 160 Hz, and thus we can be reasonably confident that temporal aliasing artifacts of the display system did not contaminate the results.

Next, we investigated the c-WWI effect with computer-generated first-order and second-order drifting gratings. One motivation for using drifting one-dimensional gratings instead of rotating wheels is that they are more amenable to a quantitative description in the Fourier domain. This is useful for any attempt to account for the illusion within a computational model, as will be seen in a later section. First-order motion is defined as a drifting modulation of pattern luminance, whereas second-order motion is defined as a drifting modulation of pattern contrast with average luminance being statistically comparable throughout the pattern. These two types of motion are known to be processed preferentially by independent hierarchically organized systems (Lu & Sperling 1995b; Seiffert & Cavanagh 1998). The so-called first-order motion system can only extract motion from first-order, but not from second-order, stimuli, whereas the "second-order" system is able to compute motion information from both types of stimuli. We found that the illusion was present and equally strong for both types of motion gratings (Fig. 29.1(c)–(d)). Even though the system's limits for motion perception were attained at much lower temporal frequencies for second-order

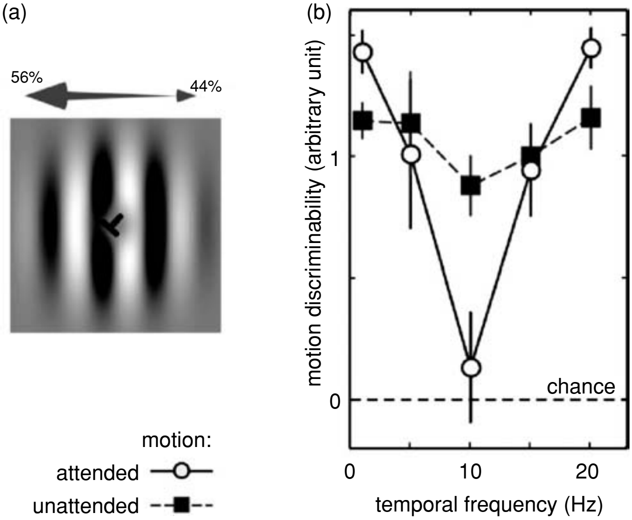


Fig. 29.2 The c-WWI and attention. When observers are required to report the direction of motion of a drifting sinusoidal grating made somewhat ambiguous by superimposing two opposite gratings of different contrasts, as in (a), their performance is specifically impaired for temporal frequencies around 10 Hz (open circles in (b). Due to temporal aliasing at this rate, the motion stimulus is perceived in the wrong direction roughly half of the time. When the observers are instructed to simultaneously attend to a sequence of rapidly changing letters at the center of the moving grating (see (a)), their pattern of response changes dramatically (black squares in (b)). The temporal aliasing effect at 10 Hz all but vanishes, implying that it is not motion processing per se, but rather the *attentional sampling* of motion information, that may rely on a quasi-periodic process.

gratings (Fig. 29.1(d)) than for first-order gratings (Fig. 29.1(c)), in both cases the peak of the illusory effect was situated around 10 Hz. This implies that the c-WWI is not simply a low-level phenomenon, attributable to peculiar properties of the first-order motion system, but is a more general characteristic of motion perception.

To summarize, under various controlled conditions the c-WWI is a genuine bistable phenomenon that occurs most strongly when the temporal frequency of the motion stimulus is around 10 Hz (VanRullen, Reddy, et al. 2005). This evidence is compatible with, and in some cases extends similar findings by, several other groups (Purves et al. 1996; Kline et al. 2004; Simpson et al. 2005).

29.3 Attentional influence

Can this illusory percept affect behavioral performance even in a task requiring objective judgments of the direction of a moving object? We showed that subjects were dramatically impaired (down to chance level) at performing a simple “left–right” motion discrimination task on drifting gratings around 10 Hz (Fig. 29.2) as long as the gratings were slightly ambiguous in terms of their motion direction (we achieved this by superimposing two opposite gratings of different contrasts). With a stimulus whose direction was easy to identify at 1 Hz or at 20 Hz, behavioral performance fell down to chance level at 10 Hz (Fig. 29.2). We suggest that temporal aliasing causes the actual stimulus direction to be wrongly perceived approximately half of the time. We also observed this effect regardless of the spatial frequency of the stimulus, as would be expected from a temporal aliasing illusion.

Using a very similar paradigm, we investigated whether the c-WWI effect depended on the availability of attentional resources. We instructed our subjects to attend to a stream of rapidly changing (every 125 msec), randomly rotated single letters presented at the center of the motion stimulus, in order to detect an L and respond to it by pressing a computer key. At the same time, as before, the subjects had to indicate the direction of the moving stimulus. Under these conditions, the motion direction discrimination performance was slightly reduced at both low (less than 5 Hz) and high (more than 15 Hz) temporal frequencies, indicating that the central letter task was effective at capturing attention. However surprisingly, with motion at 10 Hz subjects could *better* report the direction of the moving stimulus, even though they were not attending to it (Fig. 29.2). In other words, the specific impairment due to temporal aliasing at frequencies around 10 Hz was all but gone when the moving stimulus was not attended. Thus, only the processes that rely on attentional resources to compute motion direction (Cavanagh 1992; Lu & Sperling 1995a, 1995b) were affected by the temporal aliasing illusion; or simply put, attentional capture of motion information may be a quasi-periodic process.

29.4 Fourier-based motion energy model

Once temporal aliasing is demonstrated to occur for a given visual task (e.g., our motion direction judgment task), what can we conclude about the rate of the postulated underlying quasi-periodic neuronal mechanisms? If aliasing occurs at 10 Hz, does it mean that these sampling processes also function in this regime? In fact the answer is slightly more complicated (Fig. 29.3). If a periodic stimulus travels its entire cycle during one sampling interval, it will appear to be stationary, that is, with no evidence for either traveling direction. Similarly, if the stimulus travels only half of a cycle during a sampling interval, both directions of motion are equally likely interpretations; but halfway between these situations (e.g., when the stimulus travels three-fourths of a cycle during each sampling interval), then the evidence in favor of the "wrong" direction of motion will outweigh the evidence for veridical motion. Fourier motion energy analysis indicates that the maximum evidence for illusory motion is obtained exactly at three-fourths of the stimulus period. Thus, the predicted rate of subsampling for the system is always 4/3 of the observed maximum aliasing frequency. In our case, the maximal c-WWI effect at 10 Hz would be explained by discrete sampling of motion information at a rate of 13.3 Hz.

This is also what can be derived from formal Fourier analysis of motion stimuli in the spatiotemporal domain (Fig. 29.4). Using a Fourier motion energy model of motion perception (Adelson & Bergen 1985; van Santen & Sperling 1985), we predicted that the c-WWI

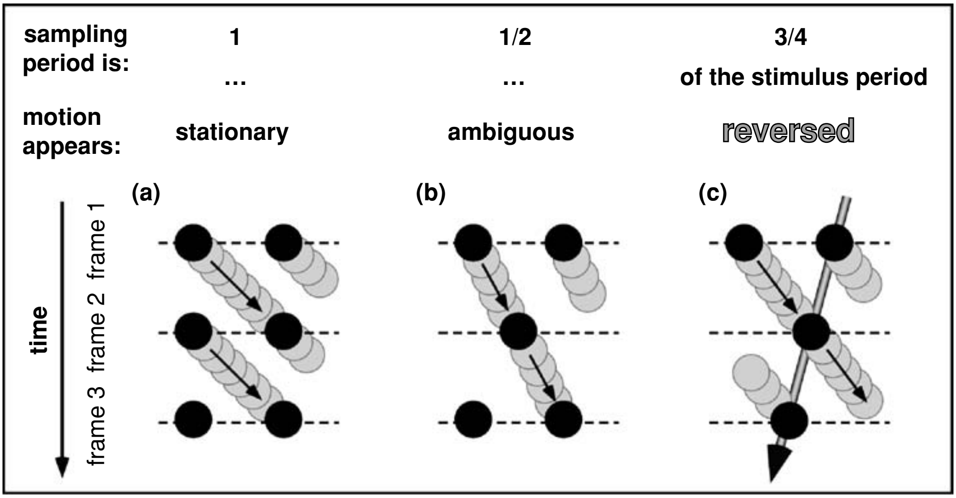


Fig. 29.3 A periodic motion stimulus is presented to a "perceptual system" that samples information periodically. If the system's sampling period is equal to the motion period (a), the stimulus will appear stationary across time. If the system's sampling period is half of the motion period, the direction of motion will appear purely ambiguous, that is, equally compatible with both alternatives (b). However, when the system's sampling period is three-fourths of the motion period, the evidence for the erroneous motion direction will be maximal and outweigh the evidence for the actual direction (c). An illusory reversed motion will be perceived. The sampling rate of a discrete system is, thus, equal to 4/3 of the temporal frequency of the stimulus that generates the maximal aliasing illusion.

would occur maximally for stimuli at 10 Hz if attentional capture of motion information was performed at a rate centered around 13.3 Hz. With a few simple assumptions of a low-pass sensitivity response to motion at higher temporal frequencies and a 50% multiplicative factor representing the average strength of attentional capture, this first model[1] of the c-WWI quantitatively accounted for all of our psychophysical measurements on one-dimensional gratings (VanRullen, Reddy, et al. 2005).

29.5 Electrophysiological correlates

To verify the plausibility of this model, we explored the electrophysiological correlates of the continuous Wagon Wheel Illusion (VanRullen et al. 2006; see also Kline et al. 2005). We contrasted the power spectrum of the EEG of twelve subjects while they experienced either the real or illusory (reversed) direction of motion with the physical stimulus being identical in both situations. Between 2 and 100 Hz, the only part of the EEG power

---

[1] In the original model described in VanRullen et al. (2005), for simplicity we had rounded up the model's sampling frequency value to 15 Hz. However, the precise value derived from Fourier analysis was 13.3 Hz, and given our more recent electrophysiological results described in Section 29.5, we now deem it important to use this value in our model. The simulations shown in Fig. 29.4 have, thus, been recalculated with an average sampling frequency of 13.3 Hz (the results are quantitatively comparable to our previous report).

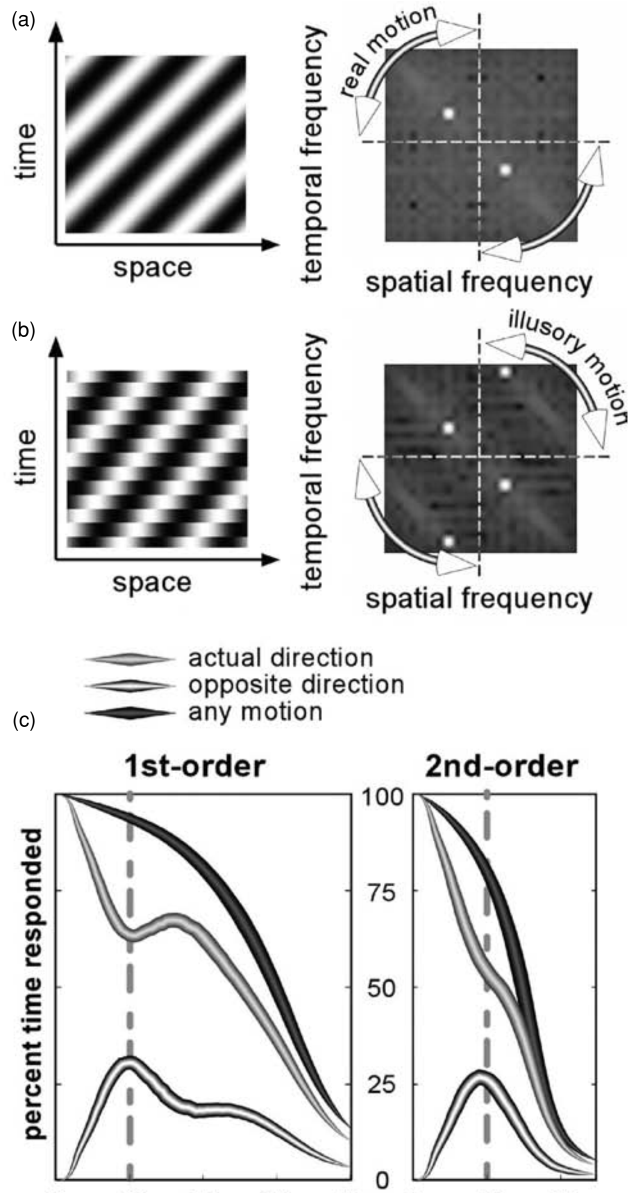


Fig. 29.4 Motion energy and temporal subsampling. (a) A one-dimensional horizontally drifting sinusoidal grating as shown in the space–time plot on the left is represented in the Fourier domain by a pair of spatiotemporal components, in diagonally opposite quadrants (right). (b) When the same motion stimulus is subsampled in time, spurious motion components appear in the Fourier spectrum. We estimate the intensity of motion perception in the actual (respectively, opposite) direction as the sum of motion energies over the corresponding quadrants of the Fourier spectrum (marked in the illustration as "real" and "illusory" motion, respectively). (c) By assuming that this simple subsampling mechanism, at an average rate of 13.3 Hz, contributes half of the total motion percept, and using low-pass motion sensitivity envelopes derived from experimental data, the model predicts rates of illusory reversals for first-order and second-order motion resembling those observed experimentally (compare with Fig. 29.1(c) and 29.1(d)).

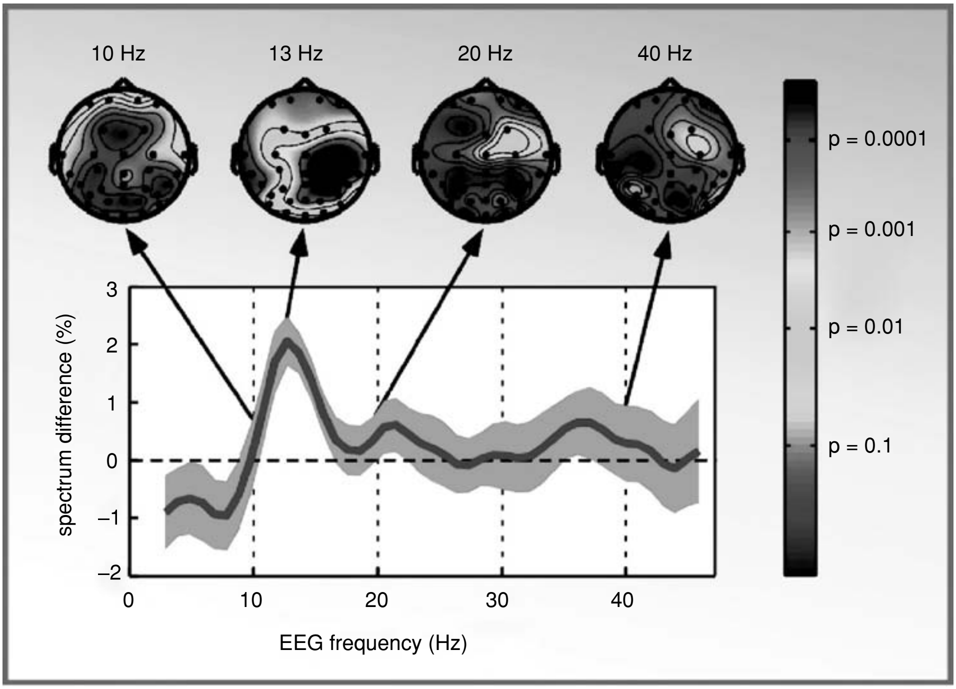


Fig. 29.5 Electrophysiological correlates of the c-WWI. We compared the power spectrum of the EEG during periods when subjects experienced real or illusory motion. A single component of the power spectrum differed significantly between the two conditions. This component was centered around 13 Hz, exactly as predicted by our model of the c-WWI based on quasi-periodic attentional sampling of motion information.

spectrum that significantly correlated with the occurrence of the c-WWI was a very specific component, centered around 13 Hz (Fig. 29.5). This component was present at exactly the same frequency when the temporal frequency of the moving stimulus was 10 Hz or 7.5 Hz. Thus, the 13-Hz effect was not driven by the stimulus itself, but rather reflected a potential "internal generator" of the c-WWI. There was no effect of perceived direction in common frequency bands of the EEG such as the theta (4–8 Hz), alpha (8–12 Hz), beta (15–25 Hz), or gamma (30–80 Hz) bands. The 13-Hz amplitude was lower during the illusion than during veridical motion perception. Similar *desynchronizations* of EEG activity (by definition, a decrease of EEG power is generally called "desynchronization") between 5 and 20 Hz have been experimentally linked to *increased* cortical activation (van Winsum et al. 1984; Pfurtscheller & Lopes da Silva 1999; Worden et al. 2000) and *increased* BOLD responses (Singh et al. 2002), which suggests that some specific neuronal source may be selectively activated during the illusion. The localization of these effects on the scalp hints at an involvement of right parietal regions. Beyond a general contribution to attentional processes, these cortical regions are known to be necessary for discriminating apparent motion and attending to the direction (onset vs. offset) of temporal events (Battelli et al.

2001; Battelli et al. 2003). Thus, both in terms of frequency specificity and topographical localization, this entire pattern of electrophysiological results is directly compatible with the predictions of our model of the c-WWI, based on quasi-periodic attentional sampling of motion information at a rate around 13 Hz.

As a complementary analysis, we attempted to use the amplitude of the 13-Hz EEG signal to predict, on a trial-by-trial basis, the direction of perceptual transitions: 13-Hz power decreased steadily in the last 2 sec before an observer reported perceiving illusory motion, and increased during the last 2 sec before the percept went back to the veridical motion. Using this relation, we could predict above chance (between 55% and 61% correct depending on the decoding strategy), on a trial-by-trial basis, the direction of upcoming transitions *before* they were actually reported by the subject. This result, which is comparable in magnitude to the quality of predictions made from monkey single-unit recordings (Britten et al. 1996), constitutes one of the first successful attempts to relate EEG activity and visual perception on a trial-by-trial basis, under conditions where perception varies while the stimulus is kept constant. This again underlines the very tight relation that seems to exist between our 13-Hz component and this particular temporal aliasing illusion.

29.6 Discussion

In this chapter we have closed a loop between psychophysical measurements, computational explanation and prediction, and electrophysiological validation. These results are compatible with the idea that attention-based motion information is computed quasi-periodically at an average rate of 13 Hz. In the following, we first elaborate on our hypothesis and its implications before addressing potential alternative accounts of the c-WWI effect.

Directing attention to the moving stimulus is necessary for the c-WWI to occur, as shown in our psychophysical experiments (VanRullen, Reddy, et al. 2005). The fact that the 13-Hz electrophysiological correlates of illusory motion perception were observed mostly over right parietal regions is important in this respect because these regions have been implicated often in various visual tasks involving selective visual attention (Buchel et al. 1998; Coull & Frith 1998; Rees & Lavie 2001; Corbetta & Shulman 2002). The same regions also display increased activity at the time of perceptual switches during bistable stimulation (Lumer et al. 1998), although for other bistable percepts such as binocular rivalry it is not clear whether the critical frequency band of the EEG is around 13 Hz as found here (Kobayashi et al. 1996) or around 40 Hz (Doesburg et al. 2005). Although patients with right parietal lesions can discriminate continuous motion and also perceive flicker normally, they tend to show bilateral deficits in the perception of apparent motion (Battelli et al. 2001) and in the discrimination of the direction (onset vs. offset) of temporal events (Battelli et al. 2003). These temporal discrimination abilities are likely to be critical for perceiving reversed motion in the c-WWI effect. This line of reasoning would thus bring us to predict that such patients may never experience the continuous Wagon Wheel Illusion. This prediction may be challenging to test directly because neglect or extinction patients generally show large-spread deficits in various visual tasks. Thus, a relative inability to judge motion direction at

10 Hz may not mean much if the patients are also performing worse than normal subjects at all temporal frequencies. However, using counterphase gratings as described in Section 29.3, our theory would predict that the patients' performance would approach the "motion unattended" curve in Fig. 29.2, that is, their performance at 10 Hz would be *better* than that of normal observers under the same conditions. Clearly, this could not be explained by some generalized visual deficit.

The prominent role of attention (and the postulated involvement of the right parietal cortex) in the illusion is also consistent with a body of data that has been presented against the hypothesis of discrete sampling of motion information. In a clever experimental manipulation, Kline et al. (2004) presented observers with two wheels rotating in opposite directions (one being the mirror image of the other). Under these conditions, subjects tended to report reversals of perceived direction (i.e., illusory motion) in only one of the stimuli at a time, even when both stimuli were placed in the same hemifield. This argued against the possibility that the visual system samples motion information in a discrete manner simultaneously over the entire visual field. In a more recent development, these authors superimposed two moving objects (one rotating, the other expanding) at the same spatial location. Again, observers generally reported perceiving illusory motion in only one object at a time (Kline et al. 2006). This was taken to argue against a spatially localized discrete sampling process. Our hypothesis does not contradict the results of these experiments. If attentional capture of motion information is really responsible for temporal aliasing in the c-WWI, then it is logical that the effects of this aliasing (i.e., the perceived reversed motion) should be visible only on the current object of our attention, be it a spatial location, one of two superimposed objects, or even a particular feature or property. Attention, indeed, can be directed to features (Treue & Martinez Trujillo 1999; Saenz et al. 2002), or objects (Baylis & Driver 1993; O'Craven et al. 1999) as well as spatial locations (Posner et al. 1980). Thus, our present findings, together with those of Kline et al. (2004, 2006), concur to support the conclusion that the c-WWI effect is limited to the – not necessarily spatial – focus of attention. In line with this hypothesis, it has been demonstrated that the Wagon Wheel Illusion affects two stimuli independently when they are perceived as separate, but simultaneously when they are perceived as a single object (VanRullen 2006).

If attention captures motion information at a rate of approximately 13 Hz, it may seem odd that motion perception in attentionally defined stimuli is limited to temporal frequencies below 5–6 Hz (Cavanagh 1992; Lu & Sperling 1995a). On the other hand, according to our hypothesis, purely attentionally defined motion stimuli at frequencies above 6.5 Hz (i.e., the critical Nyquist frequency in our model) would *always* be perceived in the wrong direction. The visual system may thus have implemented palliative strategies for counteracting this problem with the easiest one being simply to not bother processing these stimuli. This would explain the low-pass sensitivity of the attentional motion processing system.

What other explanation, besides temporally discrete sampling, could account for the c-WWI effect? Proponents of a continuous explanation have argued that elaborated Reichardt motion detectors (Reichardt 1961; Adelson & Bergen 1985; van Santen & Sperling 1985) tuned to a particular direction of motion could also experience a spurious response to the opposite direction at particular frequencies (Schouten 1967; Kline et al. 2004, 2006; Holcombe et al. 2005). Because the preferred temporal frequency of the first-order motion processing system is generally found to lie between 5 Hz and 15 Hz (Pantle 1978; Anderson & Burr 1985; Wright & Johnston 1985; Snowden & Hess 1992), it would not be wholly unexpected that reverse motion perception would peak around 10 Hz. This argument alone, however, would not predict a strong effect of attention on the illusion, because it relies on the activity of motion detectors that are largely preattentive. In addition, because the preferred sensitivity of the visual system is markedly lower for second-order stimuli (Shorter & Patterson 2001; Hutchinson & Ledgeway 2006), this hypothesis would suggest that the c-WWI effect would peak at a lower frequency for second-order than for first-order stimuli; this prediction is not supported by our data (Fig. 29.1). Finally, there does not seem to be any "a priori" reason for this explanation to predict that the illusion should be accompanied by a massive reduction of 13 Hz EEG power over right parietal regions. Overall, spurious activation of elaborated Reichardt detectors may thus be suggestive of a solution, but falls short of accounting for the complexity of the c-WWI effect.

Another, not necessarily incompatible line of thought that may explain the illusion without the need for assuming discrete subsampling relies on adaptation of neural responses to motion direction. Because the motion processing system is organized in an opponent manner, the fact that neural responses to a given motion direction decrease over time due to adaptation directly implies that the signals representing the opposite direction of motion are gaining relative strength. This is the classical explanation of the well-known motion aftereffect or "waterfall illusion." After prolonged viewing of one motion direction, a static pattern appears to move in the opposite direction. This hypothesis alone, however, would probably not be sufficient to explain the c-WWI effect. Opponent coding exists for many other visual properties such as orientation or color that also enjoy their own aftereffect following prolonged viewing, yet spontaneous reversals of perceived color while the colored stimulus remains present on the retina (i.e., an equivalent of the c-WWI effect in the color dimension) have never been reported.

Even if motion adaptation alone is unlikely to explain the illusion, it may very well contribute to it by inducing a switch in the balance between two opposite directions of motion that would already be represented in the visual system. Spurious activation of Reichardt motion detectors or discrete temporal subsampling of motion information may generate signals encoding the "wrong" direction of motion that may be too weak to dominate perception, unless they were aided by neural adaptation. Because the effects of adaptation are increased by focusing attention on the moving stimulus (Chaudhuri 1990; Shulman 1993; Lankheet & Verstraten 1995; Rees et al. 1997; Georgiades & Harris 2000, 2002; Rezec et al. 2004), even a low-level, Reichardt-based explanation, when coupled with this adaptation hypothesis, would be consistent with our findings of attentional modulation of the c-WWI effect. Yet this account would still be at a loss to explain the preserved selectivity of the illusion for second-order motion, as well as our 13-Hz EEG power correlates.

Overall, we believe that the most parsimonious explanation of the continuous Wagon Wheel Illusion is that (i) when attending a suitable moving stimulus, discrete subsampling around 13 Hz produces, through temporal aliasing, neural signals that represent the wrong direction of motion (Fig. 29.4), which are reflected in reliable changes of 13-Hz EEG power (Fig. 29.5); (ii) motion adaptation eventually helps these signals dominate perception, until they are themselves reduced by adaptation, and so on. This explains, for example, why the illusion is rarely present at stimulus onset, and generally takes a few seconds to develop. This influence of adaptation does not make the continuous Wagon Wheel effect any less interesting. In fact, there is good evidence that perceptual alternations for *any* bistable stimulus depend on some form of local adaptation (Blake et al. 2003). In this context, the fact that focal attention is necessary for the c-WWI can be interpreted in two ways: (i) Either attention contributes to discrete temporal subsampling, which is the hypothesis that we have favored in our model, and is supported by the localization of the 13-Hz EEG effect over right parietal regions; or (ii) Attention increases the effect of adaptation that would otherwise be too weak to trigger perceptual switches, with the result that the actual direction of motion would remain dominant indefinitely. The latter hypothesis appears unlikely, however, because a longer adaptation period could easily overcome the lack of attention and trigger a perceptual switch to the illusory motion. The c-WWI would be delayed, but present nonetheless in the absence of attention. In addition, the attentional effect shown in Fig. 29.2 was demonstrated under conditions where the motion direction of the stimulus was *physically reversed* every 3–4 sec on average, a situation that minimizes the influence of direction-specific adaptation. In addition, more recent experiments show that certain stimulus manipulations (e.g., change of eccentricity or contrast) can have opposing effects on the strength of the c-WWI and that of motion adaptation (VanRullen 2007). Thus, we endorse our original hypothesis, that attentional capture of motion information around 13 Hz is, so far, the most plausible candidate to account for the c-WWI effect.

If this hypothesis is, ultimately, shown to be valid for the c-WWI, it will raise the question of whether other visual domains and, perhaps, nonvisual modalities are processed in a discrete, frame-like manner or whether this is a peculiarity of motion perception. Periodic sampling of the outside environment is a ubiquitous property of sensory systems. Saccades in vision, sniffs in olfaction, whisker movements in rat somato-sensation, and even echolocation in bats or electro-location in the electric fish are all examples of explicit cyclic mechanisms for overt perceptual sampling (Uchida et al. 2006). Attention might have evolved from these periodic processes as a more economical means of covertly sampling endogenous representations, possibly relying on the widespread oscillations that can be observed in most of these systems (VanRullen et al. 2005). Temporal aliasing effects such as the c-WWI might merely be the outside manifestation of such internal periodic operations.

Acknowledgments

This research was supported by the CNRS and the NSF-ERC. We would like to thank Caitlin Berry for her considerable help with pilot experiments.

References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *J Opt Soc Am A* **2**(2): 284–299.

Anderson, S. J., & Burr, D. C. (1985). Spatial and temporal selectivity of the human motion detection system. *Vision Res* **25**(8): 1147–1154.

Battelli, L., Cavanagh, P., Intriligator, J., Tramo, M. J., Henaff, M. A., Michel, F., et al. (2001). Unilateral right parietal damage leads to bilateral deficit for high-level motion. *Neuron* **32**(6): 985–995.

Battelli, L., Cavanagh, P., Martini, P., & Barton, J. J. (2003). Bilateral deficits of transient visual attention in right parietal patients. *Brain* **126**(Pt. 10): 2164–2174.

Baylis, G. C., & Driver, J. (1993). Visual attention and objects: evidence for hierarchical coding of location. *J Exp Psychol Hum Percept Perform* **19**(3): 451–470.

Blake, R., Sobel, K. V., & Gilroy, L. A. (2003). Visual motion retards alternations between conflicting perceptual interpretations. *Neuron* **39**(5): 869–878.

Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S., & Movshon, J. A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis Neurosci* **13**(1): 87–100.

Buchel, C., Josephs, O., Rees, G., Turner, R., Frith, C. D., & Friston, K. J. (1998). The functional anatomy of attention to visual motion. A functional MRI study. *Brain* **121** (Pt 7): 1281–1294.

Burr, D. C., Ross, J., & Morrone, M. C. (1986). Smooth and sampled motion. *Vision Res* **26**(4): 643–652.

Cavanagh, P. (1992). Attention-based motion perception. *Science* **257**(5076): 1563–1565.

Chaudhuri, A. (1990). Modulation of the motion aftereffect by selective attention. *Nature* **344**(6261): 60–62.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci* **3**(3): 201–215.

Coull, J. T., & Frith, C. D. (1998). Differential activation of right superior parietal cortex and intraparietal sulcus by spatial and nonspatial attention. *Neuroimage* **8**(2): 176–187.

Crick, F., & Koch, C. (2003). A framework for consciousness. *Nat Neurosci* **6**(2): 119–126.

Doesburg, S. M., Kitajo, K., & Ward, L. M. (2005). Increased gamma-band synchrony precedes switching of conscious perceptual objects in binocular rivalry. *Neuroreport* **16**(11): 1139–1142.

Georgiades, M. S., & Harris, J. P. (2000). Attentional diversion during adaptation affects the velocity as well as the duration of motion after-effects. *Proc R Soc Lond B Biol Sci* **267**(1461): 2559–2565.

Georgiades, M. S., & Harris, J. P. (2002). Evidence for spatio-temporal selectivity in attentional modulation of the motion aftereffect. *Spat Vis* **16**(1): 21–31.

Harter, M. R. (1967). Excitability cycles and cortical scanning: a review of two hypotheses of central intermittency in perception. *Psychol Bull* **68**(1): 47–58.

Holcombe, A. O., Clifford, C. W., Eagleman, D. M., & Pakarian, P. (2005). Illusory motion reversal in tune with motion detectors. *Trends Cogn Sci* **9**(12): 559–560.

Hutchinson, C. V., & Ledgeway, T. (2006). Sensitivity to spatial and temporal modulations of first-order and second-order motion. *Vision Res* **46**(3): 324–335.

James, W. (1890). *The Principles of Psychology* (Vol. I). New York: Holt.

Kline, K., Holcombe, A. O., & Eagleman, D. M. (2004). Illusory motion reversal is caused by rivalry, not by perceptual snapshots of the visual field. *Vision Res* **44**(23): 2653–2658.

Kline, K., Holcombe, A. O., & Eagleman, D. M. (2005). Illusory motion reversal is not caused by discrete sampling of global or hemispheric visual fields. *Society for Neuroscience* 2005 abstracts #619.15.

Kline, K., Holcombe, A. O., & Eagleman, D. M. (2006). Illusory motion reversal does not imply discrete processing: reply to Rojas et al. *Vision Res* **46**(6–7): 1158–1159.

Kobayashi, T., Kato, K., Owada, T., & Kuriki, S. (1996). Difference of EEG spectral powers observed between binocular rivalry and binocular fusion. *Front Med Biol Eng* **7**(1): 11–19.

Lankheet, M. J., & Verstraten, F. A. (1995). Attentional modulation of adaptation to two-component transparent motion. *Vision Res* **35**(10): 1401–1412.

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 Res* **35**(19): 2697–2722.

Lumer, E. D., Friston, K. J., & Rees, G. (1998). Neural correlates of perceptual rivalry in the human brain. *Science* **280**(5371): 1930–1934.

O'Craven, K. M., Downing, P. E., & Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature* **401**(6753): 584–587.

Pantle, A. J. (1978). Temporal frequency response characteristic of motion channels measured with three different psychophysical techniques. *Percept Psychophys* **24**(3): 285–294.

Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin Neurophysiol* **110**(11): 1842–1857.

Pitts, W., & McCulloch, W. S. (1947). How we know universals: the perception of auditory and visual forms. *Bull Math Biophys* **9**: 127–147.

Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *J Exp Psychol Gen* **109**: 160–174.

Purves, D., Paydarfar, J. A., & Andrews, T. J. (1996). The wagon wheel illusion in movies and reality. *Proc Natl Acad Sci U S A* **93**(8): 3693–3697.

Rees, G., Frith, C. D., & Lavie, N. (1997). Modulating irrelevant motion perception by varying attentional load in an unrelated task. *Science* **278**(5343): 1616–1619.

Rees, G., & Lavie, N. (2001). What can functional imaging reveal about the role of attention in visual awareness? *Neuropsychologia* **39**(12): 1343–1353.

Reichardt, W. (1961). Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In W. A. Rosenblith (ed.), *Sensory Communication* (303–317). Cambridge, MA: MIT Press.

Rezec, A., Krekelberg, B., & Dobkins, K. R. (2004). Attention enhances adaptability: evidence from motion adaptation experiments. *Vision Res* **44**(26): 3035–3044.

Saenz, M., Buracas, G. T., & Boynton, G. M. (2002). Global effects of feature-based attention in human visual cortex. *Nat Neurosci* **5**(7): 631–632.

Schouten, J. F. (1967). Subjective stroboscopy and a model of visual movement detectors. In I. Wathen-Dunn (ed.), *Models for the Perception of Speech and Visual Form* (44–45). Cambridge, MA: MIT Press.

Seiffert, A. E., & Cavanagh, P. (1998). Position displacement, not velocity, is the cue to motion detection of second-order stimuli. *Vision Res* **38**(22): 3569–3582.

Shallice, T. (1964). The detection of change and the perceptual moment hypothesis. *British Journal of Statistical Psychology* **17**: 113–135.

Shorter, S., & Patterson, R. (2001). The stereoscopic (cyclopean) motion aftereffect is dependent upon the temporal frequency of adapting motion. *Vision Res* **41**(14): 1809–1816.

Shulman, G. L. (1993). Attentional effects of adaptation of rotary motion in the plane. *Perception* **22**(8): 947–961.

Simpson, W. A., Shahani, U., & Manahilov, V. (2005). Illusory percepts of moving patterns due to discrete temporal sampling. *Neurosci Lett* **375**(1): 23–27.

Singh, K. D., Barnes, G. R., Hillebrand, A., Forde, E. M., & Williams, A. L. (2002). Task-related changes in cortical synchronization are spatially coincident with the hemodynamic response. *Neuroimage* **16**(1): 103–114.

Snowden, R. J., & Hess, R. F. (1992). Temporal frequency filters in the human peripheral visual field. *Vision Res* **32**(1): 61–72.

Stroud, J. M. (1956). The fine structure of psychological time. In H. Quastler (ed.), *Information Theory in Psychology* (174–205). Chicago, IL: Free Press.

Treue, S., & Martinez Trujillo, J. C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature* **399**(6736): 575–579.

Uchida, N., Kepecs, A., & Mainen, Z. F. (2006). Seeing at a glance, smelling in a whiff: rapid forms of perceptual decision making. *Nat Rev Neurosci* **7**(6): 485–491.

VanRullen, R. (2006). The continuous Wagon Wheel Illusion is object-based. *Vision Res* **46**(24): 4091–4095.

VanRullen, R. (2007). The continuous Wagon Wheel Illusion depends on, but is not identical to neuronal adaptation. *Vision Res*, in press.

VanRullen, R., Guyonneau, R., & Thorpe, S. J. (2005). Spike times make sense. *Trends Neurosci* **28**(1): 1–4.

VanRullen, R., & Koch, C. (2003). Is perception discrete or continuous? *Trends Cogn Sci* **7**(5): 207–213.

VanRullen, R., Reddy, L., & Koch, C. (2005). Attention-driven discrete sampling of motion perception. *Proc Natl Acad Sci U S A* **102**(14): 5291–5296.

VanRullen, R., Reddy, L., & Koch, C. (2006). The continuous wagon wheel illusion is associated with changes in electroencephalogram power at approximately 13 Hz. *J Neurosci* **26**(2): 502–507.

van Santen, J. P., & Sperling, G. (1985). Elaborated Reichardt detectors. *J Opt Soc Am A* **2**(2): 300–321.

van Winsum, W., Sergeant, J., & Geuze, R. (1984). The functional significance of event-related desynchronization of alpha rhythm in attentional and activating tasks. *Electroencephalogr Clin Neurophysiol* **58**(6): 519–524.

Varela, F. J., Toro, A., John, E. R., & Schwartz, E. L. (1981). Perceptual framing and cortical alpha rhythm. *Neuropsychologia* **19**(5): 675–686.

White, C. (1963). Temporal numerosity and the psychological unit of duration. *Psychological Monographs: General and Applied* **77**(12): 1–37, Whole No. 575.

Worden, M. S., Foxe, J. J., Wang, N., & Simpson, G. V. (2000). Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. *J Neurosci* **20**(6): RC63.

Wright, M. J., & Johnston, A. (1985). Invariant tuning of motion aftereffect. *Vision Res* **25**(12): 1947–1955.

30

Priming and retouch in flash-lag and other phenomena of the streaming perceptual input

TALIS BACHMANN

Summary

When rapidly successive objects or object replicas are presented as sensory streams, a stimulus within a stream is perceptually facilitated relative to an otherwise identical stimulus not within the stream. Experiments on perceptual latency priming and flash-lag have convincingly shown this. Unfortunately, no consensus exists on what is (are) the mechanism(s) responsible for in-stream facilitation. Here, I discuss several alternative explanations: perceptual extrapolation of change in the specific properties of continuous stimulation, time-saving for target processing due to the early microgenetic/formation stages for target being completed on pretarget in-stream items, control of focused selective attention by the onsets of stimulus input, and preparation of the nonspecific perceptual retouch by the preceding nontarget input in stream for the succeeding target input in stream. Revisions are outlined to overcome the explanatory difficulties that the retouch theory has encountered in the face of new phenomena of perceptual dissociation.

30.1 Introduction

Objects that do not occur in isolation are processed differently compared to when they appear as separate entities. If we compare the visual latency of an object presented alone with the latency of its replica that is presented after another object (which is presented nearby in space and time), we see that the object that comes after having been primed by other input achieves awareness faster (Neumann 1982; Bachmann 1989; Scharlau & Neumann 2003a & b; Scharlau 2004). In a typical experiment, a visual prime stimulus is presented, followed by another stimulus that acts as a backward mask to the prime. In addition, another stimulus resembling the masking stimulus is presented at a different spatial position. Temporal intervals between these stimuli are varied and observers are asked to make a temporal order judgment (TOJ) regarding the stimuli. Even though the prime is backward masked up to invisibility for explicit perception, it nevertheless causes the stimulus that it precedes to appear earlier in explicit perception, compared to the simultaneous, unprimed control stimulus. This effect is called perceptual latency priming (PLP). Let us refer to the primed object as the “target” and the isolated single object as the “reference.”

Space and Time in Perception and Action, eds. Romi Nijhawan and Beena Khurana. Published by Cambridge University Press.

The stimulus onset asynchrony (SOA) between the prime and the target object that allows one to observe this effect ranges from a few dozen milliseconds (msec) to about 150–200 msec. The magnitude of the PLP effect itself usually ranges between about 20 msec and 80 msec. Various experimental controls have been used showing that the effect is not the result of response bias; not an artifact of variability in masking strength; independent of prime-mask similarity; and reflected in ERP latencies (Scharlau & Neumann 2003; Scharlau 2004a, 2007; Vibell et al. 2007).

PLP is usually studied in paired presentations of a priming stimulus and a primed target stimulus as inputs to an alert perceptual system. But there is evidence that temporal facilitation of sensory input to awareness takes place also when there are more stimulation items than merely a pair of objects. Formally, any stream that is presented to observers can be interpreted as a succession of multiple PLP displays with primes and primed items being multiply and continually presented. I define streamed presentations as a succession of stimulation items so that these items are mutually different individuated tokens and/or types. The varying retinotopic spatial position of an invariant type is interpreted as multiplication of space tokens of that type. The continuously varying sensory/perceptual quality of an invariant space token is interpreted as multiplication of perceptual types of that token. When both token and type change in discrete steps, this corresponds to a typical rapid serial visual presentation (RSVP) display.

In the flash-lag paradigm, a target item is presented within streaming input and a reference is presented out of stream. The evidence from flash-lag studies has demonstrated that when input signals that come as a sensory stream from varying spatial positions – that is, in motion (Nijhawan 1994; Whitney & Murakami 1998; Krekelberg & Lappe 1999; Brenner & Smeets 2000; Eagleman & Sejnowski 2000; Whitney et al. 2000; Kreegipuu & Allik 2003; Öğmen et al. 2004) or as a stream of sensory input from the same spatial location (Sheth et al. 2000; Bachmann & Põder 2001; Bachmann & Oja 2003) – they are used as continuous stimulation within which a target is presented, and the target's explicit perception appears to be accelerated compared to the perception of a reference that is presented in isolation. This effect is termed flash-lag effect (FLE) and, surprisingly or not, its temporal value as measured as a latency difference between perceiving the target and the reference nicely concurs with that of PLP.

Because FLE occurs both in streams with change in the feature values of the streamed sensory input (Sheth et al. 2000) and in spatially invariant streams without change in feature values of the streamed input (Bachmann & Põder 2001; Bachmann & Oja 2003), the in-stream facilitation effect appears to result from streaming per se. (In the studies with spatially and type-wise invariant streams, a spatially localized stream of overlapping **I**s was presented and a target **Z** was included in that stream. A reference **Z** was localized out of stream.) Therefore, FLE is a phenomenon that does not necessarily require a change in the spatial location of continuous streamed input nor a change in the perceptual feature attributes of the streamed input. Moreover, FLE is present with unpredictable changes in the direction of motion of the streamed stimulation (Whitney et al. 2000). It thus follows that extrapolation theories (e.g., Nijhawan 1994), assuming that stimulus items in motion are extrapolated (to compensate for the delays in neural processing and make precise catching possible) and the flashed ones are not, cannot exhaustively or universally explain the FLE: unless we accept a type of extrapolation where the spatial coordinate value of a streamed input stays invariant and its temporal value changes – a "pure" extrapolation in time.

In this chapter, I assume that PLP and FLE are the result of the workings of the same mechanism(s), whereby sensory input that is preceded in time by a spatially overlapping or adjacent other input (which comes within a critical spatiotemporal window) will be processed up to the level of explicit visibility (conscious awareness) faster than the same type of input when it is presented in isolation (flashed). I call this effect *in-stream facilitation*. Because many perceptual–attentional processing research traditions (e.g., FLE, PLP, attentional-blink, repetition-blindness versions of RSVP, varieties of visual masking, and varieties of stroboscopic motion) use streamed presentation (a fast succession of stimuli such that they are mutually different individuated tokens and/or types of stimuli) in their experimental trials, it is highly important to discover what is (are) the mechanism(s) that produce this effect. In what follows, I will present the three most likely explanatory mechanisms for in-stream facilitation and discuss their advantages and disadvantages.

30.2 Microgenetic formation of object representation in specific perceptual systems

The specific modular systems of the neocortex are embedded in a hierarchical architecture that provides the structural neuro-anatomic framework within which the active neuronal processes help create perceptual representations of the contents of environment – objects, events, scenes. These representations are built up through successive stages of perceptual formation unfolding within about 200–300 msec after stimulus presentation – a process termed "microgenesis" (Bachmann 2000). Processing of sensory input up to the stable and veridical representation passes several specific microgenetic stages about which there is a general consensus (Treisman 1998; DiLollo et al. 2000; Kanwisher 2001; Hochstein & Ahissar 2002; Crick & Koch 2003; Lamme 2003). Typically, they involve the levels of features, objects integrated from features, semantic meaning and contextual information, selective attentional control, and response selection and execution. Massive re-entrance or backpropagation of activity from higher levels more frontal in the cortex to lower levels more in the back of the cortex is well documented.

This multilevel system has to carry out several important functions. First, representing sensory qualities through feature coding. Second, integrating features into coherent objects. Third, categorizing objects. Fourth, identifying objects. (In processing input signals to become objects or in dealing with multiple objects, individuation operations are important. Objects are first registered as proto-objects or object tokens. Here we should distinguish between individuating tokens as distinct entities or individuals in time without specifying token contents as types and individuation of tokens as types having distinct specification in identity space.) Further, there are functions of *selecting* relevant objects or locations, *updating* the already active perceptual object representations while new input is continually accumulating, selecting and preparing *responses*, etc. All this takes time and it is unthinkable and computationally unrealistic that all the above mentioned functions can be carried out simultaneously and very fast. Intervals compatible with signal transmission times between only a few neurons in the range of few tens of milliseconds would be clearly insufficient. Instead, the percept unfolds microgenetically within about 150 msec, with different tasks carried out at different stages of microgenesis (Bachmann 2000; Breitmeyer & Öğmen 2006).

It should be useful to compare processing of isolated PLP or FLE reference stimuli and processing of in-stream PLP or FLE targets on the basis of this conceptualization and see how the respective processing differences within the specific cortical representational system could lead to understanding of the essence of PLP and FLE.

Note that basic features and objects integrated from these basic features are the same for flashed references and in-stream targets. Although their presentation contexts differ somewhat, the gist of their meaning is shared (and, moreover, in the PLP and FLE studies this factor has remained usually redundant). What is the difference, if any, between targets and references? An important difference comes in when we consider object updating (see also Scharlau & Neumann 2003; Moore & Enns 2004). A single-flashed reference is not updated by the following input because there is not any, but the in-stream target is followed by stream items. The latter may be interpreted either as updating of the same perceptual event with changed sensory information (in the case where the target and subsequent stream items are different – e.g., Sheth et al. 2000; Bachmann & Põder 2001) or as updating the same object file with new sensory input arriving from a new spatial location (in the case of standard FLE displays where an invariant object moves through space). The difference in the posttarget updating processes between these different types of displays would predict higher processing load in the first type of stream, smaller load in the second type of stream, and no updating load for flashed reference. Thus the value of visual latency should be smallest with single-flashed reference and largest with a target in streams of items that are different from the target. Actually, virtually all experimental results from FLE studies do not conform to this prediction.

Let us now compare the *pretarget processing* conditions (i.e., what happens before the flash) for flashed reference and in-stream target presentation. Because nothing immediately precedes the flashed reference at its location, reference-stimulus processing should begin *ab ovo* with respect to all sensory–perceptual tasks – feature processing, object integration, attentional focusing. There is nothing to update because there is only new input. The situation is different with in-stream targets. In the case when the stream represents an invariant object in motion, its processing at the features and object level has been prepared or completed already earlier (during the preceding epochs of stream presentation), and the main new task is to update spatial position. It is parsimonious to assume that this operation alone is faster than all three operations together – feature encoding, object integration, and location processing. The three operations, however, are necessary for the flashed reference stimulus to be veridically represented. Therefore, FLE is predicted as a consequence of the microgenetic preprocessing of the in-stream target–object where earlier stages of microgenesis are completed as processing the previous input (as a surrogate token for the following item) and some concomitant saving in processing time is won. Each new sample of the specific input information comes onto the foundation preprocessed already on the account of the preceding samples. These samples or tokens can be considered as proto-objects for the following input stimulation.

How does this picture conform to experimental results? Unfortunately, the data about the FLE in the so-called flash-initiated conditions where no stream occurs before flashed stimulus (i.e., a continuous stream occurs only after its presentation – Khurana & Nijhawan 1995; Eagleman & Sejnowski 2000) invalidates this explanation. In this type of experiment, no preceding information is presented for flashed references and stream-initiating targets, but the FLE is nevertheless obtained. The surrogate early microgenesis of the in-stream target object executed on the basis of the items preceding it as proto-object tokens (including the possibility where identity or type of the stream item is different from target, as in Bachmann & Põder 2001) and which saves time for in-stream conscious microgenesis of the target when it finally comes to view, cannot be used as a universal explanation of all the cases of FLE. Furthermore, if sufficiently long streams are used, FLE should increase with presenting the target at progressively later epochs of stream and stay at a stabilized level thereafter (because the stabilized and similar amount of preprocessing should be achieved at one moment). The facts showing that FLE value first increases (up to about 80 msec) with progressively later stream time points where the target is presented after stream onset and then decreases after 150 msec have passed from stream onset (remaining at a value of about 30 msec) (Bachmann & Oja 2003) contradict the "time saving in specific microgenesis" account. According to this account FLE is explained as a result of economy of time in the successive completion of percept microgenesis for the target because some of the early operations have been realized on pretarget items' processing in-stream. Thus, these stages become redundant, whereas the flashed reference processing lacks this kind of processing-time economy.

30.3 Selective spatial attention to in-stream target as the cause of in-stream facilitation

It has long been known that focusing spatial attention on the location of a stimulus decreases visual latency for that stimulus (e.g., Eriksen & Hoffman 1972; Shore et al. 2001; Scharlau 2004b). Conversely, a stimulus that appears out of the focus of attention should be processed more slowly. It has been suggested that FLE originates from the need to spend time noticing the flashed stimulus and directing attention to it (Baldo & Klein 1995; Baldo et al. 2002). Thus the faster processing of targets in stream is the consequence of selective attention being directed to the stream and not to the isolated reference stimulus out of the stream. This conceptualization can be also criticized. First, with flash-initiated displays when attention is equally distributed between the flashed reference and the to-be-moving target (that is initially located at a fixed position) and with displays where flashed and continuously moving stimuli are spatially intermingled, FLE is nevertheless obtained (Khurana

& Nijhawan 1995). Second, in the experiments where attention was explicitly controlled by precues or instructions and directed to both types of stimuli, FLE was still obtained independently from attentional selection (Khurana et al. 2000; see, however, Hommuk et al. 2008, for a bit more complex way of how attention may be involved). Moreover, FLE was uninfluenced by the cue validity suggesting that attention had its effect on shortening the RT through postperceptual processes. Third, FLE is also present when the target is presented within a stream of invariant items within the postonset epoch of stream, typically associated with attentional blink or inhibition of return (e.g., 300–500 msec after stream onset; Bachmann & Oja 2003). Although attention clearly modifies visual latencies and does this in the context of FLE displays, it cannot be used as a universal explanatory mechanism for the FLE.

30.4 Perceptual retouch theory and in-stream facilitation

Perceptual facilitation phenomena such as PLP or in-stream target benefits in FLE are expressed in terms of the relative speed with which primed target objects reach conscious representation compared to not-primed controls. That is why psychophysical adjustment and judgment procedures are typically used in order to produce these effects. If visual latency depends on how fast the stimulus inputs achieve conscious representation, it would be natural to explore how brain mechanisms responsible for upgrading preconscious sensory information into conscious representations participate in the in-stream facilitation effects. Although attention has often been regarded as responsible for consciousness of the impinging sensoriae (e.g., views of Posner, Treisman, and many others), there are many arguments against this (Baars 1997; Naccache et al. 2002; Crick & Koch 2003; Lamme 2003, 2004; Kentridge et al. 2004; Salminen et al. 2005; Koch & Tsuchiya 2007).

First, maximum concentration of undistracted spatially selective attending does not guarantee awareness of information that is in the focus of attending, such as in backward masking, binocular rivalry, or crowding effects. Second, selective attention improves processing of preconscious and subconscious information also (e.g., Carretie et al. 2005; Melcher et al. 2005; Mitroff & Scholl 2005; Montaser-Kouhsari & Rajimehr 2005) and vice versa, nonconscious cues can regulate selective attention (Jaśkowski et al. 2002; Scharlau & Ansorge 2003; Stone & Valentine 2005). Third, attention can select between different objects or stimuli that are already consciously perceived. Fourth, neurobiological structures that control attention are different from neurobiological structures that are necessary for phenomenal consciousness, although there is of course some overlap (e.g., Scheibel 1980; LaBerge 1997; Llinás & Ribary 2001). Fifth, ERP and MEG signatures of visual selective attention and visual awareness can be dissociated (Salminen et al. 2005; Schurger et al. 2008; Wyart & Tallon-Baudry 2008). Furthermore, precueing many spatial locations simultaneously so that attention is not well focused nevertheless decreases the time with which these stimuli reach awareness (Wright & Richard 2003). Thus, attention can be defined as selective activity that presets priorities irrespective of whether the *result* of this activity is also a clear and enhanced conscious apprehension of the actual sensoriae or not.

Therefore, it appears justified and timely to conceptualize a special process that is (semi)autonomous from attention whose function is to upgrade preconscious specific perceptual information into consciously experienced format. Perceptual retouch theory (Bachmann 1984, 1994, 1997) was developed exactly to achieve such an aim. As a corollary, it has been used to explain in-stream facilitation phenomena (Bachmann 1999; Bachmann et al. 2003; Bachmann & Sikka 2005). Before presenting this explanation, it is necessary to describe the theory and its premises.

There is a generally accepted distinction between two brain systems that are the most promising candidates for where to look for neural correlates of consciousness: the sensory systems for stimulus-specific phenomenal *content* (features, objects, events, and scenes) and the systems for providing necessary level of consciousness as such. The latter provides an *enabling factor* that is required for awareness but does not directly contain specific contents of conscious experiences (Bachmann 1984; Baars 1995; Bogen 1995; Llinás & Ribary 2001; Rees et al. 2002). The neurons of the content-specific system most often belong to the category termed "drivers," and the neurons of the conscious state systems mostly belong to the category termed "modulators" (Crick & Koch 1998, 2003; Sherman & Guillery 1998). The so-called nonspecific subcortical system (the NSP system) for enabling perception to manifest at the conscious level comprises first of all the extended reticular-formation activating network and some thalamic nuclei considered nonspecific for representational content (e.g., intralaminar nuclei). This system is considered the principal system for modulating cortical activity up to the level (or mode) enabling conscious awareness of the environmental stimuli and the self (Bachmann 1994; Bogen 1995; Steriade et al. 1997; John 2005; Ribary 2005; Steriade & McCarley 2005).

There is a well-documented facilitation and/or modulation of specific cortical neurons (the feature-tuned neurons that have small receptive fields and that respond to stimuli with very short delay) by the facilitatory input from the content-free "modulators" of the nonspecific thalamus (Magoun 1958; Purpura 1970; Brooks & Jung 1973; Bachmann 1984, 1994, 1997; Crick 1984; Steriade et al. 1990, 1997; Steriade 1996). Some of the thalamic structures (e.g., lateral geniculate body) belong to the relay system that transmits input from specialized sensory receptors to the cortical driver neurons. The drivers encode specific stimulus features such as size, orientation, color, and motion. Other parts of the thalamic structures termed "nonspecific" (e.g., intralaminar nuclei, nucleus reticularis thalami, and possibly pulvinar) do not participate directly in the operations of encoding of the contents of specific sensory information. Although their efferents are projected onto specific cortical driver neurons (characteristically to the apical dendrites of the middle-layer pyramidal neurons), they do not send signals about the specific contents of some sensory event. They modulate the level of activity of the specific drivers. The specific cells receive modulation not only as evoked by the input signals specific to them but also from other sources, including the input sent to some other specific stimuli whose nonspecific receptive field component is broadly overlapping but whose specific receptive field component can be different. The signal-to-noise ratio and/or gain of the activity of cortical driver units that signal the presence of some sensory feature (or combination of features) is altered by this modulation. In earlier work, I have termed this "nonspecific" system as the perceptual retouch (Bachmann 1984) or pertentional (Bachmann 1999) system, or nonspecific modulation system (NSP) in order to refer to its nonspecific neurobiological nature where its imminent lack of stimulus-encoding capability is concerned.

A newly appearing stimulus is serviced by two processes: (1) fast stimulus-specific responses by drivers and higher-level specific nodes fed by drivers, and (2) a slower, spatially dispersed but coarsely localized modulation via the collaterals through NSP. Because the latency of the maximum cortical response to a stimulus transient through nonspecific modulation is about 50–100 msec slower as compared to the afferent latency of the specific cortical neurons in response to the initial input coming through the lateral geniculate nucleus, the driver neurons, initially activated only by the specific afference, have to wait for the arrival of the stimulus-related modulatory input. (Similarly, driver neurons may have to wait for the stimulus-related *change* in the modulatory activity. This can either be the "boost" of modulation by increasing depolarization of the driver neurons' membrane potential and their firing frequency and/or resetting the phase of the synchronizing activity from NSP. The resulting effect from resetting is that oscillations in the firing activity of the cortical neurons will be set in phase with oscillations of the NSP neurons.) Importantly, this secondary input has been shown to be necessary for explicit perception (awareness) of the stimulus information pertaining initially only to preconscious specific representations carried by drivers (Bachmann 1984, 1994; Baars 1995; Bogen 1995; Newman 1995; LaBerge 1997; Steriade et al. 1997; Llinás & Ribary 2001; John 2005; Ribary 2005). Processing stimulus signals exclusively within cortical specific modular systems is insufficient for conscious perception of the data represented by these modular systems.

The evidence for the necessary effect from subcortical NSP sources projected onto specific cortical neurons in order to produce consciousness is multiple (see Bachmann 1994, and Rose 2006 – pp. 201–204, 248–257): (1) Direct stimulation of neural structures in NSP modulates awareness states and is capable of forwarding excitation to the cortex to recruit specific unit activity, which produces artificial sensations and perceptions. (2) Selectively localized injuries of NSP deprive subjects of consciousness; this is in spite of primary specific input to the cortex remaining intact. (3) The only clear-cut common site of effect of many different anesthetic substances that eliminates awareness of the environment is NSP thalamus. (4) The first sign of waking from sleep and becoming aware of surroundings is activation of NSP; partly these signs are present during REM sleep dreaming episodes when fragmentary and endogenous sentience is present. (5) When subjects are not conscious (in sleep, in pericomatose states, in anesthesia), their brains respond to external stimulation by the early, fast components of EP, but the later EP components associated with conscious perception are absent. (6) Pretrial artificial stimulation of subcortical NSP by microelectrodes in patients modifies typical backward-masking functions so that the first presented brief stimulus (usually masked by the second stimulus) becomes unusually dominant in explicit perception as if its upgrading to conscious status has been prepared ahead in time. (7) Damage of NSP or other means of inhibition of its activity leads to loss of connectivity and coherence of cortical global neuronal activity, the latter being typically the signature of conscious states. (8) In animal studies, synchronous gamma-range oscillations of cortical sensory neurons' activity that correlate with obviously conscious discrimination of the stimuli that drive these neurons is modulated by impulses from NSP. (9) Neurotransmitters that are associated with awareness, arousal, and attention such as acetylcholine and noradrenaline are released substantially by the afferents from NSP as a result of activity of subcortical neurons that are nonspecific to perceptual stimulus content. (10) When subcortical NSP of anesthetized animals is stimulated, EEG desynchronization pattern similar to the waking state is induced. (11) The number of active visual cortical neurons that respond to their specific retino-cortical afferents increases when proper sites of NSP in subcortex are stimulated. (12) Subjective aspects of perception such as timing of visual experiences and level of target visibility in masking conditions better correlate with the late, NSP-related components of ERP compared to earlier, specific driving afference-related components. In recovery from coma, emergence of first fragments of consciousness is accompanied by the emergence of late ERP components. (13) Depolarization of apical dendrites of the cortical pyramidal nerve cells caused by presynaptic influence from NSP has the dynamic pattern of decay quite similar to the decay of subjective visible persistence. (14) Damage of the nonspecific thalamus at par with intact cortical sensory areas produces disturbances in visual conscious-level perception. (15) Temporal discrimination thresholds of visual stimuli in awareness correlate with minimum intercycle intervals in the oscillatory activity of NSP units.

The whole body of facts listed above provides converging evidence for the necessarily decisive role NSP modulation has in upgrading specific activity in cortex to the level associated with sensory–perceptual awareness.

In terms of retouch theory (Bachmann 1984, 1994, 1997), the temporally delayed and spatially diffuse modulation from the NSP targeted at the specific cortical neurons that carry preconscious information about the specific stimulation content serves to "retouch" that content for visual awareness (explicit representation). It is an open question whether it is sufficient to exert just a feedforward type of (delayed) modulation onto driver neurons to obtain awareness of the driver contents, or is the repetitive re-entrant shipment of activity from the specific cortical neurons back to NSP neurons (and up again) necessary. For our present purposes, the retouch effects through NSP can be implemented in either way or both together.

Initially, at the onset of stimulation where there has not been any locally preceding visual input, such as with flashed reference or the beginning of the stream, the modulation through NSP (or the change in the dynamic characteristics of this modulation) takes considerable time to become effective. Consequently, the latency of sampling of the specific signals for explicit representation is slow (e.g., 150 msec). With the passage of time, the already presented input has had progressively more time for initiating and setting the effective nonspecific modulation. Each new set of specific signals sent via the fast afferent pathways to the cortical driver units that carry stimulus-specific information (including precise spatial position) will therefore be upgraded for explicit representation with progressively smaller visual latency until 150 msec from onset have passed. There is acceleration of the speed of conscious perception online with accumulating perceptual stimulation. After the time interval that is typical for the duration of the whole effective cycle of specific plus modulatory afferent activity has passed (i.e., about 100–200 msec), the acceleration stage is over. From now on, the information-sampling speed for explicit representation stays at a more or less stable level, but higher than it was initially (e.g., about 80–100 msec now). This explains why a moving target within stream comes to awareness more quickly than a flashed object not within a stream.

What about the puzzling problem that the flash-lag effect is present in flash-initiated conditions? Precisely the same mechanism of NSP's interaction with specific cortical units having the same temporal and spatial parameters of action as before can be used here as well. In these conditions, as ever, it takes considerable time until the boost in modulation (the wave of "retouch") will have had its effect to upgrade the stimulus-specific information up to visibility. Because NSP neurons have large receptive fields, the target of retouch (Sn), if retouch cycle has been initiated by S1 earlier, can be spatially and featurally different from S1, but nevertheless benefit from the retouch cycle initiated by S1. At the moment when the first wave of this boost of modulation arrives at specific cortical sites, the contents of specific representation have changed in the case of a moving (or otherwise featurally changing) target stimulus. The specific object is represented at its new spatial coordinates (or by some other changed feature value[s]). Specific preconscious visual processing is not slowed and conveys specific environmental information fast, but consciousness-related activity is still on the "waiting, soon to be serviced list." (Also, the decaying trace that carries previous coordinates of the target object is actively inhibited by a well-known process in the specific sensory system that helps to erase the trailing edge of the moving stimulus signals, the process typically implemented by Reichardt (1961) motion detectors. The newly arriving input, however, is not inhibited.) For the stationary reference flash, just the sensory trace featuring unchanged previous spatial coordinates (or unchanged previous feature value of the nonmoving reference-stimulus' token) is upgraded by modulation at a later moment in time. (See also Krekelberg 2001, about robust sensory persistence of position.) The feature values of the streamed stimulation and flashed stimulation will be different at the moment when the delayed nonspecific modulation arrives so as to "retouch" the contents of the specific feature level for consciousness. A change in the feature representation during the delay of retouch is not a problem because NSP's receptive fields are relatively large and cover several of the changing stimuli instances.

The attractive property of the retouch model is that it is capable of explaining some effects that many other models cannot. For example, the existence of the FLE in the conditions where stream items do not move in space and do not change feature values in feature space (Bachmann & Põder 2001) cannot be explained by the extrapolation model (Nijhawan 1994), the model proposing latency differences for moving and stationary stimuli (motion detected faster; e.g., Whitney et al. 2000), the positional averaging/integration model (Krekelberg & Lappe 1999), or the resetting positional averaging and postdiction model (Eagleman & Sejnowski 2000). However, the retouch model nicely explains this because modulation set by stream onset and the subsequent new stream items is delayed in time but becomes applied to specific signals with ever-shorter delays between the arrival of these specific signals to cortex; the arrival of impulses of the "boost" of NSP modulation does not require motion or position change (see Bachmann et al. 2003). It applies equally well to the situation where stimulus is in random motion (Murakami 2001). Perceptual retouch is automatic and spatially dispersed around its epicenter without any "intelligent" predictive capability. It also accommodates intermodal FLE and prior entry in intermodal PLP (Alais & Burr 2003; Zampini et al. 2005) because the NSP source in the retouch model is unspecific to stimulation contents and modality, reacting primarily to environmental novelty of sensory input and to coarse, general spatial direction (Bachmann 1994). Similarly, Llinás and his associates and Tononi and Edelman accept that the consciousness-generating system builds up an intermodal, integrated consciousness, absorbing the workings of the specific representational neurons into the synchronized activity of the generalized modulator, which itself is unspecific in terms of contents.

In recent work, some important findings about the timing of specific cortical processing in the context of visual latencies and FLE have come to be known (Arrighi et al. 2005; McDonald et al. 2005). It appears that latencies of specific cortical responses explain neither FLE nor audiovisual TOJ under attentional prior entry conditions. What matters is not neural response onset latencies, but (1) the amplitude of the cortical response to a transient, and (2) neural integration times. This is encourageing, because in the perceptual retouch model what counts is how fast the specific responses are integrated for consciousness and the strength of the facilitatory modulation from NSP that has its effect only later, after the first components of the specific cortical response characterized by its latency according to the first arriving input are already over (Bachmann 1984, 1994). This means that not the early specific cortical activity, but the later, modulated activity is associated with explicitly perceptible contents of stimulation. Just this type of information forms the basis of psychophysical estimations and judgments that exemplify PLP and FLE. Moreover, if NSP impulses that modulate specific activity need a certain minimum time for execution of effective integration through oscillating synchrony, then latencies of the *initial* cortical input – from the specific system or NSP alike – are not suitable predictors of timing of awareness.

Yet, the retouch theory has its difficulties. In some experimental setups of the FLE displays the flashed object is presented within the continuously moving object, but FLE is nevertheless obtained (e.g., Nijhawan & Khurana 2000; Nijhawan 2001). Particularly problematic is when a stationary disc is flashed centered exactly within the moving annulus, but perceived as displaced out of it (creating an illusory crescent-like image formed from the inner edge of the annulus on the side in the direction of motion and the outer edge of the flashed disc, which is apparently half-visible within the annulus as if lagging behind it in space). The retouch theory in its original form meets difficulty in explaining this because the flashed object is presented from within the area that is within an object that has already received preprocessing during its preceding exposure before the flash. Therefore, perceptual retouch from NSP should have been prepared well ahead in time and the flashed disc should not be seen lagging behind the continuously moving annulus.

Another related problem emerges from findings (e.g., Cai & Schlag 2001) concerning the illusory spatial dissociation of different properties of a gradually changing object. Thus, when an object in motion changes its size and at one particular instant of size change its color also changes, perceivers misalign the locations of shape and color change so that a wrong color will be misbound with a wrong-sized object from the later instances of the presented shape. (Similarly, Moutoussis and Zeki 1997, showed illusory dissociation of color and motion direction information in awareness when a set of colored squares moving back and forth changed motion direction at about 3 Hz frequency and synchronously with color change. For the change of these two features to appear subjectively simultaneously, color had to change 60–100 msec later than motion direction.) Due to the nonspecific nature of presynaptic modulation of the specific neurons, delayed perceptual retouch should upgrade the whole specific set of object attributes as they were in actuality, although from a later instance in space and time. In order to keep the explanatory viability of perceptual retouch theory, it has been modified to accommodate these feature-dissociation effects.

30.5 The revised perceptual retouch account

The above-mentioned difficulties for the perceptual retouch theory of in-stream perception (and concomitant perceptual illusions) are in one way or another related to some effects of dissociation in representing the various specific perceptual information pertaining to the one and the same, physically integrated, object. In other words, why is it that the putative modulation sent from NSP to the cortex “finds” the object-specific information in a dissociated state? A disarmingly simple answer would be that retouch impulses arrive when specific cortical processing of the different specific properties of the changing object has reached different stages of completion. Although the size and location of the changing shape at the moment of color-changing flash in Cai and Schlag (2001) experiment may be registered fast, the change in color may take longer to register, and therefore the new value of the color of the shape will be misbound to the new-sized shape that was presented later than the shape where the actual color change took place. Analogously, the change in motion direction and color in Moutoussis and Zeki study (1997) may take different time to register in the specific system, and the “traveling moment” of the retouch activity finds dissociated specific features.

There are two possibilities, not necessarily mutually exclusive, why change in color may take longer to register. First, color processing may be generally slower compared to motion processing. Second, in the described experiments, motion and shape are attributes with continuously changing spatial values, but color change is discontinuous and therefore feeds the NSP system less persistently. In light of the preceding assumptions it is obvious that nonspecific modulation in the model should be made more closely related to the feature-binding functions. In its original form (Bachmann 1984, 1997) it was not.

Fortunately, there are thorough studies in which the consciousness-generating and feature-binding functions of the nonspecific thalamus (the NSP) are made to coexist (e.g., Llinás et al. 1998; Singer 1998; Steriade 2000; Llinás & Ribary 2001; Llinás et al. 2005). The specific classical afferent pathways transmit specific environmental information and establish a gamma-frequency synchronized activity within the neuronal ensembles that stand for the actual sensory input (including objects). For this information to become integrated into a conscious perceptual image, the respective activity should be synchronized with another gamma-band synchronous oscillatory activity generated by the NSP (e.g., intralaminar nuclei). As the set of specific neurons participating in the oscillatory ensemble may be recruiting different neuronal pools (that represent different objects and/or their different features) at different epochs of time, illusory dissociations in the specific characteristics of actually coherent and integrated objects can be expected.

If at the moment when nonspecific synchronizing activity arrives, only part of the specific information pertaining to the target object has been encoded by the specific neurons, only that part will be integrated into conscious representation at that moment. Specific information that is slower in being established as a specific representation "enters" the integrated dynamic core at a later moment. Synchronization by and with NSP modulation need not be localized in time at one narrow instant. It is a continuous, temporally extended process. Thereby, dissociation in time of separate features of an object can be possible. But to obtain this, special artificial conditions have to be used where the natural speed of specific-plus-nonspecific integration through gamma-range synchrony cannot cope with the spatiotemporal properties of experimental stimulation. The objective, physical change of featured objects in space–time should violate natural brain-process constraints of spatiotemporal processing for consciousness. Visual masking, the Cai and Schlag effect, PLP, and FLE achieve this. It is important to stress that the binding problem relates not only to binding within specific cortical representational systems but also to binding for *conscious* representation. For conscious-level perceptual representation, *two* coordinated binding activities are necessary.

Hence, the revised perceptual retouch model assumes that NSP modulation is expressed not only as the increased excitatory postsynaptic potentials of the specific driver neurons and a related increase in firing frequency (and decrease in the latency with which the drivers begin to discharge; see Bachmann 1997), but that the crucial effect of this modulation consists in (re)setting the synchrony between gamma-range oscillations in the activity of the specific representational neurons and nonspecific thalamic neurons. For this purpose, temporal binding mechanisms based on local cortical circuits of coincidence detection may be used (Llinás et al. 2002). Coincidence detection by coactivation of the specific cortical and nonspecific thalamic nuclei that results in synchronized (or coherent) gamma-band frequency of the cortical pyramidal neurons (that represent selected specific perceptual information) is the essence of the mechanism of cognitive binding. (GABA-ergic mechanisms seem to be important in forming restricted but continuously changing cortical areas NSP modulated up to consciousness, with the current actual content – Llinás et al. 2005.)

Because the newly presented sensory input activates specific neurons (drivers and higher cortical nodes) sooner than the nonspecific counterpart of the oscillatory activity arrives at the cortex, the later-arriving signals from the delayed specific input are what determine the contents of perceptual awareness. When the stimulus item remains invariant, it is represented veridically and by its specific trace at the end of or after its presentation; for instance, in terms of its initial location of the unchanging feature values. When the stimulus undergoes change while the nonspecific "retouch" is being generated, then what comes to the conscious experience at first corresponds to the changed specific contents of stimulation. The new set of specific neurons representing the updated environmental stimulation will be integrated by the already functional nonspecific oscillatory system. This means that the changing stimulation (i.e., the streamed stimulation after the flash) will be represented in consciousness with a smaller latency than the unchanging reference stimulation (which essentially *is* FLE) and that presentation of a priming stimulus speeds up visual latency to awareness for the succeeding primed stimulus (PLP). Formally, this is similar to any abstract model assuming (i) "some form of neural facilitation applied along the inferred trajectory of [changing] objects" (Khurana & Nijhawan 1995, p. 566), and (ii) the absence of explicit perception for some critical time such as 100 msec (Khurana & Nijhawan 1995). From the neurobiological point of view, however, the retouch model uniquely relates "neural facilitation" and perceptual delay to the NSP processes originating from subcortical nuclei and constituting the core of the system dedicated to upgrading consciousness. Yet more "inferred motion" and other similar semi-intelligent concepts are not necessary. Furthermore, extrapolation cannot explain FLE in invariant-item streams without motion (Bachmann & Põder 2001) and in unpredictable motion change conditions (Whitney et al. 2000). NSP is nonspecific, extrapolation is specific to the vector of change. Perceptual retouch trails what has happened by default; extrapolation "foresees" future feature values.

Some new research findings are consistent with this picture. Taylor et al. (2005) showed that sustained attention to a morphing shape (monkeys were monitoring continuous change of an object they should have held in consciousness) was accompanied by the gamma-range oscillatory synchronized activity in V4. When the successive objects underwent abrupt changes, each new appearance of a different object was accompanied by a burst of synchronized gamma activity that decayed afterward. Summerfield et al. (2002) showed that if certain masked stimuli entered awareness, they induced a burst of gamma activity. Some computational models of how the background gamma rhythms in cortical local circuits may mediate attentive/conscious perception have been also developed (Börgers et al. 2005).

In some of the studies conducted by my research group, we have found that FLE is largest and release from backward masking strongest when the target is presented within the initial epoch stream of about 50–150 msec after stream onset (Bachmann & Oja 2003; Bachmann & Sikka 2005). It is logical to assume that the gamma-frequency modulation generated after stimulation onset should also be most conspicuous exactly at that time epoch. Indeed, as shown by Herrmann and Mecklinger (2001), this is the case. Taylor et al. (2005) find similarly delayed gamma bursts in response to presented stimulus items. (At the same time it should be noted that in the conditions where targets are presented in the very first epoch of a stream and less than 50 msec after its onset, backward masking of the target by stream items prevails over facilitation that emerges only later. Thus when the stream is very short and stimuli can have masking effects, in-stream facilitation may be outweighed by masking effects. This is the likely reason why in auditory streaming studies when using very short streams, in-stream decrement has been found (e.g., Bregman & Rudnicky 1975).

Yet, why is it that in experiments by Nijhawan (2001) and Nijhawan and Khurana (2000) the disc flashed within a continuously moving annulus appears spatially offset? The annulus information has been on for some time and both its feature binding in the specific system that formed the corresponding object Gestalt and NSP-binding oscillations have been completed. The only relatively simple task is to update spatial position of the already visible integrated object. This is all fast now. For the flashed reference disc, however, many operations within the specific representational system have to begin anew, and its specific microgenesis is relatively slow. Moreover, among its specific features is the cue that it is spatially invariant. Therefore, an apparent conflict between specific spatial attributes of the moving annulus and the flashed disc is represented by the retouching activity because retouch "finds" these objects at different stages of completion and with different spatial coordinates because of the relatively more accelerated conscious microgenesis for the annulus. In addition, the masking-like inhibitory effect of the annulus on the disc within the specific processing system could be considered (Kirschfeld 2006). The part of the annulus that arrives at the critical spatial location of the disc later in time should have stronger masking effect because backward masking is typically stronger than forward or simultaneous masking.

30.6 General discussion

The description of the revised perceptual-retouch model has kept the main foundations of the original version. Let me relate this general approach to the theoretical disputes persistent in the current research on microgenetic issues pertinent to space–time problems.

30.6.1 Well-known unsolved controversies in the context of the retouch model

Visible representations of static and dynamic stimulation are formed with a delay and timed according to how quickly the NSP modulation arrives at cortex. As the conscious-perception timing is anchored to NSP modulation that comes late in time, there is no mystique in Benjamin Libet's results showing long latencies of sensory awareness after stimulus application (Libet 2004). In vision, the latency should be measured by the time it takes for NSP to have caused the modulation by gamma-range synchronization of the specific active representations in cortical specific cells. This is about 150–200 msec. The controversial reports by Libet about the apparent subjective antedating of sensory stimulation in time require, however, extra commentary. In the key experiment, the cortical train of stimulation in the somatosensory area took about 500 msec to elicit conscious sensation. If during this "consciously silent" time a stimulus was applied to the skin, the subject reported that the sensation generated at the skin appeared before the cortically induced sensation. The stimulus that was presented later was perceived as being presented before the one that was onset earlier. It was also apparent that the skin stimulus elicited both the fast cortical ERP components and the slower ones. The cortical stimulus had no distinct and typical ERP responses. My interpretation is that the "normal" skin stimulus elicited the natural train of afferent impulses both through specific and nonspecific pathways and instigated the retouch activity, which is necessary for conscious sensations. Direct cortical stimulation had only indirect influence on NSP, and therefore it was more difficult to elicit fast enough conscious responses. There is no need for the mysterious antedating. Simply in one case the normal NSP modulation necessary for sensory awareness of stimulation was evoked, and in the other case this modulation was much delayed. This is why subjective latency difference was found. (Another possibility is related to the effect that less distinct sensory stimuli appear to be presented earlier than more distinct stimuli, although actually the order is just the other way around – Bachmann et al. 2004. Because in Libet's experiment the skin pulse was near threshold, its weakness may have created the temporal order illusion that it has been presented earlier than the more strongly felt skin sensation created artificially by cortical stimulation.)

Eagleman (e.g., this volume) suggests that due to the temporal variability (about 80 msec) of sensory latencies for different visual characteristics of stimuli, the brain has to wait for this time before committing to an integrated percept, otherwise no spatiotemporally coherent objects could be experienced. This idea is not at odds with the retouch account. If retouch is slow (however necessary for consciousness, as has been proved by neurobiological research), it arrives no sooner than preconscious integrations in the specific sensory areas have been completed. We could hypothesize that the gamma-range oscillations within the set of cortically specific neurons complete their job at interfeatural integration, and only then the nonspecific gamma-range oscillations will be resonated onto them. The consciousness process is late because awareness is functionally and adaptively more apt at representing a coherent and comprehensible environment and not noise dispersed in space–time.

30.6.2 Preconscious processing and the retouch model

Although NSP-mediated conscious experience comes slowly, this does not mean that specific brain areas are not processing specific sensory information in the meantime. Beautiful experiments by Mitroff and Scholl (2004, 2005) demonstrate that even if sensory objects are not represented in consciousness (for instance, by motion-induced blindness; see Bonneh et al. 2001), their specific characteristics continue to be processed preconsciously. Thus, when a bar gradually changes its orientation while out of awareness and is switched off abruptly, it re-enters conscious perception after this off-transient. Importantly, it does this with the feature value (orientation) it acquired during the preconscious time of processing and not the last value of when it was consciously perceived before disappearing from awareness. Updating functions can be well and veridically performed by the specific system preconsciously. Perceptual retouch is applied to updated, not old, specific contents even if updating itself proceeds out of awareness. But as psychophysical judgments and evaluations in PLP and FLE are based on conscious percepts, the object-updating model of FLE and PLP without invoking the NSP processes seems to be insufficient.

30.6.3 Does perceptual retouch operate like a periodic sampler?

It is possible that the NSP-based modulation for consciousness does not work steadily, but produces semidiscrete sampling in a repetitive snapshot-like manner (Crick & Koch 2003; John 2005; VanRullen, Reddy & Koch, this volume). Thus, the wagon-wheel illusion with continuous exposure can be maximized at about 10-Hz luminous stimulation. This could be a temporal measure of the frequency with which nonspecific gamma-frequency modulation is itself cyclically amplified. Consequently, perceptual-awareness functions would also vary with the same rhythm. Technically, a wavelet analysis of 10-Hz gamma bursts of neuronal cortical activity could be envisaged. Consider a combined use of 40-Hz and 10-Hz wavelets.

To conclude, the description provided on the preceding pages (1) emphasizes the need to accept that there is a specialized but "nonspecific" mechanism for perceptual consciousness apart from selective attention; (2) shows how the perceptual retouch mechanism featuring coordinated interplay between (fast-responding) specific representational neurons in cortex and (slowly responding) nonspecific modulatory systems in subcortex explains most of the experimental findings concerning FLE and PLP; (3) stresses the need to accept the revision of the old retouch model by virtue of emphasizing the role of gamma-frequency modulation, which does accomplish feature binding for consciousness and thus overcomes some of the theoretical difficulties posed by the new data about perceptual feature dissociations. Specific objects in the shape of their cortical representations are embraced by nonspecific waves of gamma rhythm, thus enabling them to be consciously seen.

Acknowledgments

This work was partly supported by Estonian Science Foundation grant #7118. The author thanks Romi Nijhawan, Beena Khurana, and anonymous reviewers for their help during preparation of this chapter.

References

Alais, D., & Burr, D. (2003). The "flash-lag" effect occurs in audition and cross-modally. *Curr Biol* **13**: 59–63.

Arrighi, R., Alais, D., & Burr, D. (2005). Neural latencies do not explain the auditory and audio-visual flash-lag effect. *Vision Res* **45**: 2917–2925.

Baars, B. (1995). Surprisingly small subcortical structures are needed for the state of waking consciousness, while cortical projection areas seem to provide perceptual contents of consciousness. *Conscious Cogn* **4**: 159–162.

Baars, B. J. (1997). *In the Theater of Consciousness: The Workspace of the Mind.* Oxford: Oxford University Press.

Bachmann, T. (1984). The process of perceptual retouch: nonspecific afferent activation dynamics in explaining visual masking. *Percept Psychophys* **35**: 69–84.

Bachmann, T. (1989). Microgenesis as traced by the transient paired-forms paradigm. *Acta Psychologica* **70**: 3–17.

Bachmann, T. (1994). *Psychophysiology of Visual Masking: The Fine Structure of Conscious Experience*. Commack, NY: Nova Science Publishers.

Bachmann, T. (1997). Visibility of brief images: The dual-process approach. *Conscious Cogn* **6**: 491–518.

Bachmann, T. (1999). Twelve spatiotemporal phenomena, and one explanation. In G. Aschersleben, T. Bachmann, & J. Müsseler (eds.), *Cognitive Contributions to the Perception of Spatial and Temporal Events*. (173–206). Amsterdam: Elsevier.

Bachmann, T. (2000). *Microgenetic Approach to the Conscious Mind.* Amsterdam/ Philadelphia: John Benjamins.

Bachmann, T., Luiga, I., Põder, E., & Kalev, K. (2003). Perceptual acceleration of objects in stream: evidence from flash-lag displays. *Conscious Cogn* **12**: 279–297.

Bachmann, T., & Oja, A. (2003). Flash-lag without change in feature space is alive and well at late intervals after stream onset. *Perception* **32**(Suppl.): 126–127.

Bachmann, T., & Põder, E. (2001). Change in feature space is not necessary for the flash-lag effect. *Vision Res* **41**: 1103–1106.

Bachmann, T., Põder, E., & Luiga, I. (2004). Illusory reversal of temporal order: the bias to report a dimmer stimulus as the first. *Vision Res* **44**: 241–246.

Bachmann, T., & Sikka, P. (2005). Perception of successive targets presented in invariant-item streams. *Acta Psychologica* **120**: 19–34.

Baldo, M. V. C., & Klein, S. A. (1995). Extrapolation or attention shift? *Nature* **378**: 565–566.

Baldo, M. V. C., Kihara, A. H., Namba, J., & Klein, S. A. (2002). Evidence for an attentional component of the perceptual misalignment between moving and flashing stimuli. *Perception* **31**: 17–30.

Bogen, J. E. (1995). On the neurophysiology of consciousness: I. An overview. *Conscious Cogn* **4**: 52–62.

Bonneh, Y., Cooperman, A., & Sagi, D. (2001). Motion-induced blindness in normal observers. *Nature* **411**: 798–801.

Börgers, C., Epstein, S., & Kopell, N. J. (2005). Background gamma rhythmicity and attention in cortical local circuits: a computational study. *Proc Natl Acad Sci U S A* **102**: 7002–7007.

Bregman, A. S., & Rudnicky, A. J. (1975). Auditory segregation: stream or streams? *J Exp Psychol Hum Percept Perform* **104**: 263–267.

Breitmeyer, B. G., & Öğmen, H. (2006). *Visual Masking: Time Slices Through Conscious and Unconscious Vision.* Oxford: Oxford University Press.

Brenner, E., & Smeets, J. B. J. (2000). Motion extrapolation is not responsible for the flash-lag effect. *Vision Res* **40**: 1645–1648.

Brooks, B., & Jung, R. (1973). Neuronal physiology of the visual cortex. In R. Jung (ed.), *Handbook of Sensory Physiology. Vol. VII/.3* (325–440). New York: Springer-Verlag.

Cai, R., & Schlag, J. (2001). A new form of illusory conjunction between color and shape [Abstract]. *J Vis* **1**: 127a.

Carretie, L., Hinojosa, J. A., Mercado, F., & Tapia, M. (2005). Cortical response to subjectively unconscious danger. *NeuroImage* **24**: 615–623.

Crick, F. (1984). Function of the thalamic reticular complex: the searchlight hypothesis. *Proc Natl Acad Sci U S A* **81**: 4586–4590.

Crick, F., & Koch, C. (1998). Constraints on cortical and thalamic projections: the no-strong-loops hypothesis. *Nature* **391**: 245–250.

Crick, F., & Koch, C. (2003). A framework for consciousness. *Nat Neurosci* **6**: 119–126.

DiLollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: the psychophysics of reentrant visual processes. *J Exp Psychol Gen* **129**: 481–507.

Eagleman, D. M. (2010). How does the timing of neural signals map onto the timing of perception? In R. Nijhawan & B. Khurana (eds.), *Space and Time in Perception and Action* (216–231). Cambridge: Cambridge University Press.

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science* **287**: 2036–2038.

Eriksen, C. W., & Hoffman, J. E. (1972). Temporal and spatial characteristics of selective encoding from visual displays. *Percep Psychophys* **12**: 201–204.

Herrmann, C. S., & Mecklinger, A. (2001). Gamma activity in human EEG is related to high-speed memory comparisons during object selective attention. *Vis Cogn* **8**: 593–608.

Hochstein, S., & Ahissar, M. (2002). View from the top: hierarchies and reverse hierarchies in the visual system. *Neuron* **36**: 791–804.

Hommuk, K., Bachmann, T., & Oja, A. (2008). Precuing an isolated stimulus temporarily outweighs in-stream stimulus facilitation. *J Gen Psychol* **135**: 167–181.

Jaśkowski, P., van der Lubbe, R., Schlotterbeck, E., & Verleger, R. (2002). Traces left on visual selective attention by stimuli that are not consciously identified. *Psychol Sci* **13**: 48–54.

John, E. R. (2005). From synchronous neuronal discharges to subjective awareness? In S. Laureys (ed.), *Progress in Brain Research.* **150**: 143–171.

Kanwisher, N. (2001). Neural events and perceptual awareness. *Cognition* **79**: 89–113.

Kentridge, R. W., Heywood, C. A., & Weiskrantz, L. (2004). Spatial attention speeds discrimination without awareness in blindsight. *Neuropsychologia* **42**: 831–835.

Khurana, B., & Nijhawan, R. (1995). Extrapolation or attention shift? *Nature* **378**: 566.

Khurana, B., Watanabe, K., & Nijhawan, R. (2000). The role of attention in motion extrapolation: are moving objects "corrected" of flashed objects attentionally delayed? *Perception* **29**: 675–692.

Kirschfeld, K. (2006). Stopping motion and the flash-lag effect. *Vision Res* **46**: 1547–1551.

Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. *Trends Cogn Sci* **11**: 16–22.

Kreegipuu, K., & Allik, J. (2003). Perceived onset time and position of a moving stimulus. *Vision Res* **43**: 1625–1635.

Krekelberg, B. (2001). The persistence of position. *Vision Res* **41**: 529–539.

Krekelberg, B., & Lappe, M. (1999). Temporal recruitment along the trajectory of moving objects and the perception of position. *Vision Res* **39**: 2669–2679.

LaBerge, D. (1997). Attention, awareness, and the triangular circuit. *Conscious Cogn* **6**: 149–181.

Lamme, V. A. F. (2003). Why visual attention and awareness are different. *Trends Cogn Sci* **7**: 12–18.

Lamme, V. A. F. (2004). Separate neural definitions of visual consciousness and visual attention; a case for phenomenal awareness. *Neural Netw* **17**: 861–872.

Libet, B. (2004). *Mind Time. The Temporal Factor in Consciousness*. Cambridge, MA: Harvard University Press.

Llinás, R., Leznik, E., & Urbano, F. J. (2002). Temporal binding via cortical coincidence detection of specific and nonspecific thalamocortical inputs: a voltage-dependent dye-imaging study in mouse brain slices. *Proc Natl Acad Sci U S A* **99**: 449–454.

Llinás, R., & Ribary, U. (2001). Consciousness and the brain. The thalamocortical dialogue in health and disease. *Ann N Y Acad Sci* **929**: 166–175.

Llinás, R., Ribary, U., Contreras, D., & Pedroarena, C. (1998). The neuronal basis for consciousness. *Philos Trans R Soc Lond B Biol Sci* **353**: 1841–1849.

Llinás, R., Urbano, F. J., Leznik, E., Ramirez, R. R., & van Merle, H. J. F. (2005). Rhythmic and dysrhythmic thalamocortical dynamics: GABA systems and the edge effect. *Trends Neurosci* **28**: 325–333.

Magoun, H. W. (1958). *The Waking Brain*. Springfield, IL: C. C. Thomas.

McDonald, J. J., Teder-Sälejärvi, W. A., Di Russo, F., & Hillyard, S. A. (2005). Neural basis of auditory-induced shifts in visual time-order perception. *Nat Neurosci* **8**: 1197–1202.

Melcher, D., Papathomas, T. V., & Vidnyansky, Z. (2005). Implicit attentional selection of bound visual features. *Neuron* **46**: 723–729.

Mitroff, S. R., & Scholl, B. J. (2004). Last but not least: seeing the disappearance of unseen objects. *Perception* **33**: 1267–1273.

Mitroff, S. R., & Scholl, B. J. (2005). Forming and updating object representations without awareness: evidence from motion-induced blindness. *Vision Res* **45**: 961–967.

Montaser-Kouhsari, L., & Rajimehr, R. (2005). Subliminal attentional modulation in crowding condition. *Vision Res* **45**: 839–844.

Moore, C. M., & Enns, J. T. (2004). Object updating and the flash-lag effect. *Psychol Sci* **15**: 866–871.

Moutoussis, K., & Zeki, S. (1997). Functional segregation and temporal hierarchy of the visual perceptive system. *Proc R Soc Lond B Biol Sci* **264**: 1407–1414.

Murakami, I. (2001). A flash-lag effect in random motion. *Vision Res* **41**: 3101–3119.

Naccache, L., Blandin, E., & Dehaene, S. (2002). Unconscious masked priming depends on temporal attention. *Psychol Sci* **13**: 416–424.

Neumann, O. (1982). Experimente zum Fehrer-Raab-Effekt und das Wetterwart-Modell der visuellen Maskierung. *Report No. 24/1982, Psychological Institute, Ruhr-University Bochum.*

Newman, J. (1995). Thalamic contributions to attention and consciousness. *Conscious Cogn* **4**: 172–193.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature* **370**: 256–257.

Nijhawan, R. (2001). The flash-lag phenomenon: object motion and eye movements. *Perception* **30**: 263–282.

Nijhawan, R., & Khurana, B. (2000). Conscious registration of continuous and discrete visual events. In T. Metzinger (ed.), *Neural Correlates of Consciousness*. (203–219). Cambridge, MA: MIT Press.

Öğmen, H., Patel, S. S., Bedell, H. E., & Camuz, K. (2004). Differential latencies and the dynamics of the position computation process of moving targets, assessed with the flash-lag effect. *Vision Res* **44**: 2109–2128.

Purpura, D. (1970). Operations and processes in thalamic and synaptically related neural subsystems. In F. O. Schmitt (ed.), *The Neurosciences. Second Study Program* (458–470). New York: Rockefeller University Press.

Rees, G., Kreiman, G., & Koch, C. (2002). Neural correlates of consciousness in humans. *Nat Rev Neurosci* **3**: 261–270.

Reichardt, W. (1961). Autocorrelation: a principle for the evaluation of sensory information by the central nervous system. In W. A. Rosenblith (ed.), *Principles of Sensory Communication*. New York: Wiley.

Ribary, U. (2005). Dynamics of thalamo-cortical network oscillations and human perception. In S. Laureys (ed.), *Progress in Brain Research, Vol.* **150** (127–142).

Rose, D. (2006). *Consciousness*. Oxford: Oxford University Press.

Salminen, N., Koivisto, M., & Revonsuo, A. (2005). Independence of visual awareness from attention at early processing stages. Paper presented at *ASSC9*, June 24–27, Pasadena, CA.

Scharlau, I. (2004a). Evidence against response bias in temporal order tasks with attention manipulation by masked prime. *Psychol Res* **68**: 224–236.

Scharlau, I. (2004b). The spatial distribution of attention in perceptual latency priming. *Q J Exp Psychol* **57**: 1411–1436.

Scharlau, I. (2007). Temporal processes in prime-mask interaction. *Advances in Cognitive Psychology* **3**: 241–255.

Scharlau, I., & Ansorge, U. (2003). Intention-mediated control of attention by nonconscious information (paper presented at *ASSC8*).

Scharlau, I., & Neumann, O. (2003a). Temporal parameters and time course of perceptual latency priming. *Acta Psychologica* **113**: 185–203.

Scharlau, I., & Neumann, O. (2003b). Perceptual latency priming by masked and unmasked stimuli: evidence for an attentional interpretation. *Psychol Res* **67**: 184–196.

Scheibel, A. (1980). Anatomical and physiological substrates of arousal: a view from the bridge. In J. A. Hobson & M. A. B. Brazier (eds.), *The Reticular Formation Revisited: Specifying Function for a Nonspecific System* (55–66). New York: Raven.

Schurger, A., Cowey, A., Cohen, J., Treisman, A., & Tallon-Baudry, C. (2008). Distinct and independent correlates of awareness and attention in a hemianopic patient. *Neuropsychologia* **46**: 2189–2197.

Sherman, S. M., & Guillery, R. W. (1998). On the actions that one nerve cell can have on another: distinguishing “drivers” from “modulators.” *Proc Natl Acad Sci U S A* **95**: 7121–7126.

Sheth, B. R., Nijhawan, R., & Shimojo, S. (2000). Changing objects lead briefly flashed ones. *Nat Neurosci* **3**: 489–495.

Shore, D. J., Spence, C., & Klein, R. M. (2001). Visual prior entry. *Psychol Sci* **12**: 205–212.

Singer, W. (1998). Consciousness and the structure of neuronal representations. *Philos Trans R Soc Lond B Biol Sci* **353**: 1829–1840.

Steriade, M. (1996). Awakening the brain. *Nature* **383**: 24–25.

Steriade, M. (2000). Corticothalamic resonance, states of vigilance and mentation. *Neurosci* **101**: 243–276.

Steriade, M., Jones, E. G., & Llinás, R. R. (1990). *Thalamic Oscillations and Signalling*. New York: Wiley.

Steriade, M., Jones, E. G., & McCormick, D. A. (1997). *Thalamus*. Oxford: Elsevier.

Steriade, M., & McCarley, R. W. (2005). *Brain Control of Wakefulness and Sleep*. New York: Springer/Kluwer/Plenum.

Stone, A., & Valentine, T. (2005). Orientation of attention to nonconsciously recognised famous faces. *Cogn Emot* **19**: 537–558.

Summerfield, C., Jack, A. I., & Burgess, A. P. (2002). Induced gamma activity is associated with conscious awareness of pattern masked nouns. *Int J Psychophysiol* **44**: 93–100.

Taylor, K., Mandon, S., Freiwald, W. A., & Kreiter, A. K. (2005). Coherent oscillatory activity in monkey area V4 predicts successful allocation of attention. *Cerebral Cortex* **15**: 1424–1437.

Treisman, A. (1998). Feature binding, attention and object perception. *Proc R Soc Lond B Biol Sci* **353**: 1295–1306.

VanRullen, R., Reddy, L., & Koch, C. (2010). A motion illusion reveals the temporally discrete nature of visual awareness. In R. Nijhawan & B. Khurana (eds.), *Space and Time in Perception and Action* (521–535). Cambridge: Cambridge University Press.

Vibell, J., Klinge, C., Zampini, M., Spence, C., & Nobre, A. C. (2007). Temporal order is coded temporally in the brain: early event-related potential latency shifts underlying prior entry in a cross-modal TOJ task. *J Cogn Neurosci* **19**: 109–120.

Whitney, D., & Murakami, I. (1998). Latency difference, not spatial extrapolation. *Nat Neurosci* **1**: 656–657.

Whitney, D., Murakami, I., & Cavanagh, P. (2000). Illusory spatial offset of a flash relative to a moving stimulus is caused by differential latencies for moving and flashed stimuli. *Vision Res* **40**: 137–149.

Wright, R. D., & Richard, D. M. (2003). Sensory mediation of stimulus-driven attentional capture in multiple-cue displays. *Percept Psychophys* **65**: 925–938.

Wyart, V., & Tallon-Baudry, C. (2008). Neural dissociation between visual awareness and spatial attention. *J Neurosci* **28**: 2667–2679.

Zampini, M., Shore, D. I., & Spence, C. (2005). Audiovisual prior entry. *Neurosci Lett* **381**: 217–222.

Index